Online ISSN : 0975-4172
Print ISSN : 0975-4350

Global Journal of Computer Science & Technology

DISCOVERING THOUGHTS AND INVENTING FUTURE

Highlights

- Web Page Recommendation
- Robust Digital Watermarking
- Sensor Data Encryption
- Analyzing Earth Movers

Laboratory, Antarctica

Volume 12 | Issue 9 | Version 1.0

ENG

Global Journal of Computer Science and Technology

Global Journal of Computer Science and Technology

Volume 12 Issue 9 (Ver. 1.0)

Open Association of Research Society

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

(A Delaware USA Incorporation with "Good Standing"; Reg. Number: 0423089)

Sponsors:Global Association of Research

Open Scientific Standards

Publisher's Headquarters office

Global Journals Inc., Headquarters Corporate Office, Cambridge Office Center, II Canal Park, Floor No. 5th, ***Cambridge (Massachusetts)***, Pin: MA 02141 United States

USA Toll Free: +001-888-839-7392

USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Association of Research, Marsh Road, Rainham, Essex, London RM13 8EU United Kingdom.

Packaging & Continental Dispatching

Global Journals, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investers@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Including by Air Parcel Charges):

For Authors:
22 USD (B/W) & 50 USD (Color)
Yearly Subscription (Personal & Institutional):
200 USD (B/W) & 250 USD (Color)

Editorial Board Members (HON.)

Dr. Bart Lambrecht
Director of Research in Accounting and FinanceProfessor of Finance
Lancaster University Management School
BA (Antwerp); MPhil, MA, PhD (Cambridge)

Dr. Carlos García Pont
Associate Professor of Marketing
IESE Business School, University of Navarra
Doctor of Philosophy (Management), Massachusetts Institute of Technology (MIT)
Master in Business Administration, IESE, University of Navarra
Degree in Industrial Engineering, Universitat Politècnica de Catalunya

Dr. Fotini Labropulu
Mathematics - Luther College
University of ReginaPh.D., M.Sc. in Mathematics
B.A. (Honors) in Mathematics
University of Windso

Dr. Lynn Lim
Reader in Business and Marketing
Roehampton University, London
BCom, PGDip, MBA (Distinction), PhD, FHEA

Dr. Mihaly Mezei
ASSOCIATE PROFESSOR
Department of Structural and Chemical Biology, Mount Sinai School of Medical Center
Ph.D., Etvs Lornd University
Postdoctoral Training,
New York University

Dr. Söhnke M. Bartram
Department of Accounting and FinanceLancaster University Management SchoolPh.D. (WHU Koblenz)
MBA/BBA (University of Saarbrücken)

Dr. Miguel Angel Ariño
Professor of Decision Sciences
IESE Business School
Barcelona, Spain (Universidad de Navarra)
CEIBS (China Europe International Business School).
Beijing, Shanghai and Shenzhen
Ph.D. in Mathematics
University of Barcelona
BA in Mathematics (Licenciatura)
University of Barcelona

Philip G. Moscoso
Technology and Operations Management
IESE Business School, University of Navarra
Ph.D in Industrial Engineering and Management, ETH Zurich
M.Sc. in Chemical Engineering, ETH Zurich

Dr. Sanjay Dixit, M.D.
Director, EP Laboratories, Philadelphia VA Medical Center
Cardiovascular Medicine - Cardiac Arrhythmia
Univ of Penn School of Medicine

Dr. Han-Xiang Deng
MD., Ph.D
Associate Professor and Research Department Division of Neuromuscular Medicine
Davee Department of Neurology and Clinical NeuroscienceNorthwestern University Feinberg School of Medicine

Dr. Pina C. Sanelli
Associate Professor of Public Health
Weill Cornell Medical College
Associate Attending Radiologist
NewYork-Presbyterian Hospital
MRI, MRA, CT, and CTA
Neuroradiology and Diagnostic Radiology
M.D., State University of New York at Buffalo,School of Medicine and Biomedical Sciences

Dr. Roberto Sanchez
Associate Professor
Department of Structural and Chemical Biology
Mount Sinai School of Medicine
Ph.D., The Rockefeller University

Dr. Wen-Yih Sun
Professor of Earth and Atmospheric SciencesPurdue University Director
National Center for Typhoon and Flooding Research, Taiwan
University Chair Professor
Department of Atmospheric Sciences, National Central University, Chung-Li, TaiwanUniversity Chair Professor
Institute of Environmental Engineering, National Chiao Tung University, Hsin-chu, Taiwan.Ph.D., MS The University of Chicago, Geophysical Sciences
BS National Taiwan University, Atmospheric Sciences
Associate Professor of Radiology

Dr. Michael R. Rudnick
M.D., FACP
Associate Professor of Medicine
Chief, Renal Electrolyte and Hypertension Division (PMC)
Penn Medicine, University of Pennsylvania
Presbyterian Medical Center, Philadelphia
Nephrology and Internal Medicine
Certified by the American Board of Internal Medicine

Dr. Bassey Benjamin Esu
B.Sc. Marketing; MBA Marketing; Ph.D Marketing
Lecturer, Department of Marketing, University of Calabar
Tourism Consultant, Cross River State Tourism Development Department
Co-ordinator , Sustainable Tourism Initiative, Calabar, Nigeria

Dr. Aziz M. Barbar, Ph.D.
IEEE Senior Member
Chairperson, Department of Computer Science
AUST - American University of Science & Technology
Alfred Naccash Avenue – Ashrafieh

Contents of the Volume

i. Copyright Notice
ii. Editorial Board Members
iii. Chief Author and Dean
iv. Table of Contents
v. From the Chief Editor's Desk
vi. Research and Review Papers

1. Web Page Recommendation Approach Using Weighted Sequential Patterns and Markov Model. *1-7*
2. Layered and Feature Based Image Segmentation Using Vector Filtering. *9-13*
3. Robust Digital Watermarking for Color Images Using Fuzzy Vault. *15-18*
4. Computer-Based Decision Support System: A Study of Akanu Ibiam Federal Polytechnic. *19-22*
5. Sensor Data Encryption Protocol for Wireless Network Security. *23-26*
6. Automated Listing for Exportable Products. *27-34*
7. Optimization and Efficient Transmission Schedule for 802.11 Networks with Jamming Characteristics. *35-40*
8. Advanced Methods to Improve Performance of K-Means Algorithm: A Review. *41-45*
9. TCP Performance Over Mobile IP Wired-cum-Wireless Networks. *47-53*

vii. Auxiliary Memberships
viii. Process of Submission of Research Paper
ix. Preferred Author Guidelines
x. Index

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Web Page Recommendation Approach Using Weighted Sequential Patterns and Markov Model

By K. Suneetha & Dr.M.Usha Rani
Sree Vidyanekethan Engineering College/JNTU

Abstract - Web page recommendation aims to predict the user's navigation through the help of web usage mining techniques. Currently, researchers focus their attention to develop a web page recommendation algorithm using the well known pattern mining techniques. Here, we have presented a web page recommendation algorithm using weighted sequential patterns and markov model. To mine the weighted sequential pattern, we have modified the prefixspan algorithm incorporating the weightage constraints such as, spending time and recent visiting. Then, the weighted sequential patterns are utilized to construct the recommendation model using the Patricia trie-based tree structure. Finally, the recommendation of the current users is done with the help of markov model that is the probability theory enabling the reasoning and computation as intractable. For experimentation, the synthetic dataset is utilized to analyze the performance of W-Prefixspan algorithm as well as web page recommendation algorithm. From the results, the memory required for the W-prefixSpan algorithm is less than 50% of memory needed for PrefixSpan algorithm.

Keywords : *Web page recommendation, Weighted sequential pattern, Prefixspan, Patricia-trie, Markov model.*

GJCST Classification: *I.5.m*

Strictly as per the compliance and regulations of:

Web Page Recommendation Approach Using Weighted Sequential Patterns and Markov Model

K. Suneetha [α] & Dr.M.Usha Rani [σ]

Abstract - Web page recommendation aims to predict the user's navigation through the help of web usage mining techniques. Currently, researchers focus their attention to develop a web page recommendation algorithm using the well known pattern mining techniques. Here, we have presented a web page recommendation algorithm using weighted sequential patterns and markov model. To mine the weighted sequential pattern, we have modified the prefixspan algorithm incorporating the weightage constraints such as, spending time and recent visiting. Then, the weighted sequential patterns are utilized to construct the recommendation model using the Patricia trie-based tree structure. Finally, the recommendation of the current users is done with the help of markov model that is the probability theory enabling the reasoning and computation as intractable. For experimentation, the synthetic dataset is utilized to analyze the performance of W-Prefixspan algorithm as well as web page recommendation algorithm. From the results, the memory required for the W-prefixSpan algorithm is less than 50% of memory needed for PrefixSpan algorithm.

Keywords : Web page recommendation, Weighted sequential pattern, Prefixspan, Patricia-trie, Markov model.

I. Introduction

Sequential pattern mining, an advance of association rule mining, is an imperative subject of data mining, often applied for extracting the useful information [9]. Sequential pattern mining algorithms deals with the problem of determining the frequent sequences in a given database [6]. Sequential pattern is a sequence of itemsets that often occur in a specific order and thus, all items in the same itemset are expected to encompass the same transaction time value or within a time gap. Each sequence have a temporally ordered list of events, wherein each event is a compilation of items (itemset) occurring simultaneously. The temporal ordering among the events is induced by the absolute timestamps associated with the events [10]. Generally, all the transactions of a customer are collectively viewed as a sequence, called customer - sequence, where each transaction is represented as an itemset in that sequence and all the transactions are scheduled in a particular order based on the transaction-time [8]. Recently, there has been a substantial interest in using sequential mining approaches to construct web page recommendation systems.

Here, we have presented a web page recommendation algorithm using weighted sequential patterns and markov model. The overall process of page recommendation based on Web usage mining consists of three phases: data preparation, mining of weighted sequential patterns and recommendation.

Data preparation: This step consists of, (i) identifying the different users' sessions from the usually very poor information available in web log files and (ii) reconstructing the users' navigation path within the identified sessions.

Mining of weighted sequential patterns: The traditional sequential pattern mining problem is extended by allowing a weight to be associated with each page in a user session to reflect interest of each page within the user session. In turn, this provides with an opportunity to associate a weight parameter with each page in a resulting sequential pattern, which called a weighted sequential pattern (WSP).

Recommendation: After mining the weighted sequential patterns, the patricia-based tree is constructed. From the Patricia tree, a recommendation model is developed based on markov model for predictions of users to find web pages they want to visit.

II. Related Work

Literature presents a lot of web page recommendation algorithms based on web used mining, collaborative filtering, and rule-based filtering. Here, we portray some recent works presented in the literature for web page recommendation.

Forsati, R *et al* [3] have developed an effective and scalable approach to deal with the web page recommendation problem. Here, a distributed learning machine has been employed to learn the performance of previous users' and to cluster the pages based on learned pattern. Dealing with unvisited or recently added pages was one of the difficult and challenging tasks in recommendation systems. As they would never be

Author α : Sr. Lecturer Dept. of MCA, SVEC, Tirupati Andhrapradesh, INDIA-517102. E-mail: umasuni.k@gmail.com
Author σ : Assoc.Professor & HEAD, Dept. of Computer Science, SPMVV, Tirupati, Andhrapradesh, INDIA- 517102.
E-mail: musha_rohan@yahoo.com

recommended, it is indispensable to provide a chance for these seldom visited or recently added pages to be incorporated in the recommendation set. By considering this problem, a Weighted Association Rule mining algorithm has been presented for the recommendation purposes. Also, a HITS algorithm has been exploited to extend the recommendation set. Furthermore, they have analyzed the proposed algorithm under various settings, and revealed the efficiency of this approach in enhancing the overall quality of web recommendations.

Yicen Liu *et al* [15] have introduced an automatic tag recommendation algorithm that has been employed in the large-scale and real-time data process successfully and efficiently. Most of the prior researches on tag suggestion have focused on initially, discovering the relationship between testing and training data and then, assigning the top ranked tags of the most related training data to the testing object. But, they not paid any attention in determining the internal relationship between the tags and weblogs. In their current research, more than 43% of tags, which have been labeled by weblog users, have really been employed in the body of the text. In the meantime, the term frequency distribution, the paragraph frequency distribution, and the first occurrence position of tags were dissimilar from the ones of non-tags in the text. As well, the tags of a weblog have been assigned in two steps. Initially, some probability distributions of the word attributes have been trained through the labeled training weblogs, and some keywords of a testing weblog have been extracted as one part of the tags based on the probability distributions. Subsequently, with the aid of Latent Semantic Indexing (LSI) model, the other parts of the tags have been obtained from the first part ones. Experiments conducted on an extensive tagging dataset of weblogs 12 have confirmed that the average tagging time for a new weblog was less than 0.02 sec, and more than 74% of testing weblogs have been properly labeled by means of the top 15 tags.

An approach for recommendations of unvisited pages has been presented by Forsati, R *et al* [11]. They have focused on the recommender systems based on the user's navigational patterns and provided proper recommendations to cater to the current needs of the user. The group of users with analogous browsing patterns has been identified by employing an offline data preprocessing and clustering technique. The experiments conducted on real usage data from a commercial web site have demonstrated a considerable enhancement in the recommendation efficiency of the proposed system.

Web Personalization is viewed as an application of data mining and machine learning approaches to create models of user behavior that can be applied to the task of forecasting the user needs and adapting future interactions with the eventual goal of enhanced user satisfaction. An extensive overview of intelligent methods for Web Personalization has been presented by Sarabjot Singh Anand and Bamshad Mobasher [12]. They have studied the state-of-the-art in Web personalization. Initially, a depiction of the personalization process and a classification of the current techniques to Web personalization have been presented. Also, they have discussed the different sources of data available to personalization systems, the modeling techniques utilized, and the current techniques to analyze these systems. Numerous challenges faced by the researchers in developing these systems and also the solutions to these challenges proposed in literature have been described. They have concluded with a discussion on the open challenges that must be addressed by the research community if this technology is to create a positive impact on user satisfaction with the Web.

Due to the increasing number of Web sites such as e-businesses contain a huge number of pages, users find it very hard to swiftly reach their own target pages. Thus, Hiroshi Ishikawa *et al* [5] have proposed two approaches to Web usage mining as a key solution to these problems. First of all, an efficient recommendation system called the L-R system has been depicted, which creates user models through classifying the Web access logs and by mining access patterns based on the transition probability of page accesses, and then, recommend the significant pages to the users based on both the user models and the Web contents. The prototype system has been analyzed and obtained the positive effects. Secondly, another approach has been employed for creating user models that clusters the Web access logs based on the access patterns. Moreover, the user models assist to find the unexpected access paths corresponding to ill-formed Web site design. In addition, Daniel Mican and Nicolae Tomai [2] have proposed WRS, architecture for robust web applications. Within the structure, usage data was being implicitly obtained by data collection sub-module. Here, the usage data has been extracted, online and in real time, via a proactive technique. They have efficiently exploited association rule mining among both frequent and infrequent items for successful pattern discovery. This was due to the fact that the pattern discovery module transactionally processes users' sessions and employs incremental storage of rules. Also, they have proved that the Wise Recommender System (WRS) has been straightforwardly implemented within any web application, because of the efficient integration of the three phases into an online transactional process.

III. Proposed Algorithm to Web Page Recommendation Based on Weighted Sequential Access Patterns

Web page recommendation is significant research over the past decade due to its real world application. With the intention of real world applicability, we have developed an approach for web page recommendation using weighted sequential pattern and markov model. Here, the traditional sequential pattern mining algorithm is modified significantly by incorporating the significant measure to mine more useful patterns. Then, the markov model described in [4] is used to recommend the web pages. The block diagram of the proposed approach for web page recommendation is given in Figure 1. The important steps for generating recommendations to the user is as follows,

- Data Preprocessing
- W-PrefixSpan for mining of weighted sequential web access patterns
- Building Pattern tree model
- Generation of recommendations using markov model

Block diagram

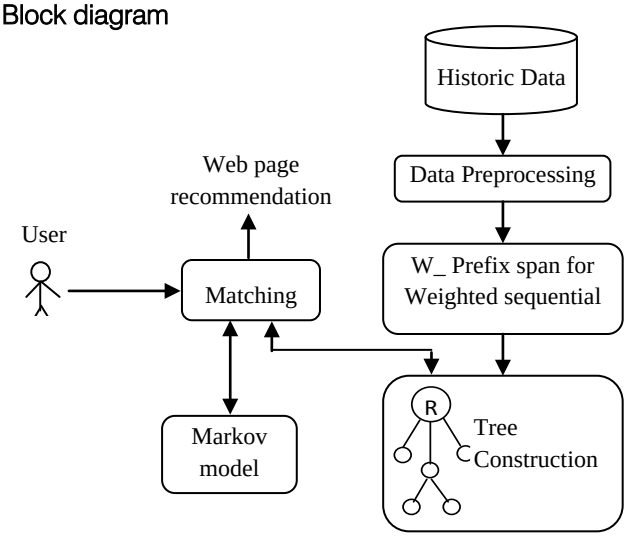


Fig. 1: Block diagram of the proposed approach to web page recommendation

a) Data Preprocessing

This section describes the preprocessing steps of web log file that is the input of the proposed web page recommendation approach. In general, the web log file consists of, IP address, access time, HTTP request method used, URL of the referring page and browser name (for an example, *Web server log file: 192.162.37.21 [23/Feb/2012:08:17:25]"GET / HTTP/1.1" "http://www.sigkdd.org/kddcup/index.php" IE/7.0 Windows 07*). The initial process of the proposed web page recommendation approach is to preprocess the web log file such a way that the mining process should be applied. Here, we make use of the sequential pattern mining process so there is a need to convert the web log file into the sequential database that should be in the proper format to mine the weighted sequential patterns.

User Identification: User identification is an important step for constructing the sequential database. IP address and user session are utilized here to track a unique user from the web log file. Unique IP address is a new user but at the same time, the user session should be fixed for particular time period. If the user session is reached to a particular duration for the same IP address, then the new session is acted like new user. Based on this, user transaction is formed from the web log file.

Weighted Sequential database generation: Once the user transactions are identified, the weighted sequential database is generated including the sequence of web pages visited by the user, time spent by the user on corresponding web page and its recent information.

Let assume the web log database D having IP address, access time, HTTP request method used, URL of the referring page and browser name. After applying the data preprocessing steps, we generate the weighted sequential database that is represented as follows, W_{ij} in which, 'i' belongs to the set of users and 'j' signifies the set of pages visited by the corresponding user. Here, every element of w_{ij} contains three tuples $w_{ij} = (p_{ij}, s_{ij}, r_{ij})$ in which the first tuple belongs to the web pages, the second one belongs to the time spent within that page and the third one belong to whether its recent one or not. For example, $w_{ij} = (p_1, 20, 0)$ is a tuple, which denotes that the user spent 20 seconds on page p_1 and '0' signifies that the page p_1 is not recently accessed by the user.

b) W-Prefixspan for Mining of Weighted Sequential Web Access Pattern

Once the weighted sequential database is constructed, the mining procedure is carried out to find the interesting sequential patterns. Here, PrefixSpan [7] is modified as W-PrefixSpan (Weighted-PrefixSpan) by incorporating the ***spending time*** and ***recent view*** into the mining procedure. The weightage measure assumed in the proposed W-PrefixSpan algorithm is spending time and recent view. The two aspects taken for providing the weightage of the sequential patterns are,

Spending time: One of the fields in the *web log data* is the duration of the web page which is viewed by the user. Generally, time spent by the user within a particular page is necessary to identify the importance of web pages. From the web page which having long duration, we can conclude that this particular web page has been referred by the user in a long occasion

because of its worth. Thus, the *spending time* is an important measure for the researchers who are attempting to identify the interest of the users. So, if we incorporate the time duration into the mining of sequential patterns, the interesting relationships can be found out from the mined sequential patterns that can be effectively applied to web page recommendation process.

Recent view: Another significant measure taken for sequential pattern mining is recent view that describes whether the page is accessed recently or not. The reason behind taking the recent view for mining the sequential pattern is that the more importance should be given for the web pages which are accessed recently than the older one. The behavior of the user surely vary depend on the time so the recent behavior of the user is significant for finding the sequence analysis. With the intention of behavior variation over time, the recent view is also incorporated into the sequential pattern mining algorithm to achieve a subset of more interesting sequential patterns (SR-Patterns).

W-PrefixSpan algorithm: In this section, we describe an efficient algorithm, W-PrefixSpan, for mining all the SR- patterns from weighted sequence databases. The W-PrefixSpan algorithm is developed by modifying the eminent PrefixSpan algorithm, which uses the pattern growth methodology for mining the frequent sequential patterns recursively. Let $W_{ij} = \langle (p_{11}, s_{11}, r_{11}), (p_{12}, s_{12}, r_{12}), \cdots, (p_{1n}, s_{1n}, r_{1n}) \rangle$ be a data sequence of weighted database W_{ij}, where p_{ij} is web page, s_{ij} is a spending time and r_{ij} signifies the recent view. A sequence $W_s = \langle (p_{11}, s_{11}, r_{11}), (p_{12}, s_{12}, r_{12}), \cdots, (p_{1m}, s_{1m}, r_{1m}) \rangle$ is said to be a sub sequence of W_{ij} only if, (1) W_s is a subsequence of W_{ij}, $W_{ij} \in W_s$

(2) $t_1 < t_2 < \cdots < t_m$ where, t_1 is the time at which p_{ij} occurred in W_s, $1 \le r \le m$. A sequence is said to be SR sequence W_{ij} if and only if, (1) W_s is a subsequence of W_{ij}, (2) the W-support should be satisfied.

At first, the weighted sequential database W_{ij} is given to the proposed W-PrefixSpan algorithm that discovers the 1-length weighted sequential patterns from the weighted sequential database by scanning the database once. The 1-SR patterns (spending time with recent view) which satisfy the predefined support threshold are mined from the sequential database by simply scanning the database. The W-support for the 1-length pattern is computed as follows,

$$W\_\sup(p) = \frac{1}{N} \frac{\sum_{i=1}^{N_T \in p} I_s(i) * R(i)}{\sum_{i=1}^{N_T \in p} R(i)}$$

Where, N → Number of user transaction in the weighted sequential database N_T → Number of transaction that contains the web page p, $R(i)$ → recent information

$$I_s(i) = \sum_{i=1}^{N_T} \left(\frac{s_i}{\sum_{i=1}^{M_T} s_i} \right)$$

Where, M_T → Total number of web pages in one transaction, s_i → spending time Then, the projection database is formed by projecting the collection of postfixes of mined 1-SR sequence. In projection database, 'n' disjoint subsets are generated if the mined 1-SR patterns contain 'n' number of sequence. Then, the 2-length SR-patterns are mined from the projected database by computing the weighted support on the projected database. Again, the projected database is formed with the help of mined 2-SR patterns and this process is repeated recursively until all SR sequential patterns are mined. The following provides a detailed explanation of the important steps involved in the proposed W-PrefixSpan algorithm.

Input : A weighted sequence database W_{ij}, and the minimum support threshold $\min\_W\_\sup$.

Output : The complete set of weighted-sequential patterns β.

Method : Call *W-PrefixSpan*$(\langle\ \rangle, 0, W_{ij})$.

Subroutine : *W-PrefixSpan* ($\alpha, l, W_{ij}|_\alpha$)

Parameters: α : SR-sequential pattern; l : the length of α; $W_{ij}|_\alpha$: the α-projected database, if $\alpha \neq \langle\ \rangle$; otherwise, the weighted sequence database W_{ij}.

Method:

1. Scan $W_{ij}|_\alpha$ once, find the set of SR- pattern s such that
   a) 's' can be assembled to the last element of α to form a SR-sequential pattern; or
   b) $\langle s \rangle$ can be appended to α to form a SR-sequential pattern.
2. For each SR- web page s, append it to α to form a SR-sequential pattern α'.
5. For each α', construct α'-projected database $W_{ij}|_\alpha$, and call *PrefixSpan* ($\alpha', l+1, W_{ij}|_\alpha$).

c) Building of Pattern Tree Model

Once we mine the weighted sequential patterns, the pattern tree is constructed using the procedure defined in [14, 13]. Initially, trie-based data structure given in [1] is used to construct the pattern tree for web page recommendation. Later, the modification was done by [14, 13], who utilized the patricia-based data structure for web page recommendation due to the advantages of particia structure over the trie structure. Here, the procedure defined in [14, 13] is applied to the proposed approach for constructing tree structure. The method for constructing the pattern tree in the proposed web page recommendation approach is as follows: 1) Generate an empty root node, 2) Add the most sub pattern in the SR-sequential pattern set into a node next to the root node, 3) Insert the postfixes of pattern into child node only if the current pattern to be inserted is a super pattern of inserted patterns, 4) Otherwise, current pattern is inserted into the node next to the root node, and 5) Step 3 and step 4 is repeated for every pattern in the mined SR-pattern set.

d) Generation of Recommendations Using Markov Model

This section describes the markov model utilized for web page recommendation. Here, we make use of the markov model described in [4] that is used in the identication of the next page to be accessed by the Web site user based on the sequence of previously accessed pages. Here, whenever a new user comes to get the recommendation, the sequence path of the new user is matched with the patricia-trie structure. Then, the subsequent web page whether it may be from same node or from its child node is retrieved. Now, the sequence path of the new user is used to find the accurate recommendation using the probability definition used in the previous work [4].

Let the input sequence visited by the user be $s_1, s_2, \ldots, s_n$. At first, the sequence, $s_1, s_2, \ldots, s_n$ is matched with the patricia-trie structure that can provide the matched result like, $s_1, s_2, s_n, s_{n+1}, \ldots, s_k$, $s_1, s_2, s_n, s_{n+1}, \ldots, s_l$ and $s_1, s_2, s_n, s_{n+1}, \ldots, s_m$. Every sequence has the weighted support value that is given in the node of constructed tree. Then, the probability of computation is carried out to find the most important sequence for the user. The following equation provides the probability of estimated value for the matched result.

$$s_{n+1}^{(1)} = \{P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_k) \mid s_1, s_2, \ldots, s_n)\}$$

$$s_{n+1}^{(2)} = \{P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_l) \mid s_1, s_2, \ldots, s_n)\}$$

$$s_{n+1}^{(3)} = \{P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_m) \mid s_1, s_2, \ldots, s_n)\}$$

This probability, $pro(s_{n+1} \mid s)$, is estimated by using all sequences of all users in tree structure constructed from the weighted sequential database W_{ij}.

$$P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_m) \mid s_1, s_2, \ldots, s_n) = \frac{W\_\sup(s_1, s_2, \ldots, s_n)}{W\_\sup(s_1, s_2, s_n, s_{n+1}, \ldots, s_m)}$$

Then, the final recommendation is based on the following equation:

$$s_{n+1} = \arg\max\{s_{n+1}^{(1)}, s_{n+1}^{(2)}, s_{n+1}^{(3)}\}$$

IV. Results and Discussions

This section presents the results obtained from the experimentation and its detailed discussion about the results. The proposed approach of web page recommendation is experimented with the synthetic dataset and the result is evaluated with the precision, applicability and hit ratio.

a) Experimental Set Up and Dataset Description

The proposed web page recommendation approach is implemented in Java (jdk 1.6). Here, the synthetic dataset is generated as like the same format of real datasets and the performance of the proposed approach is evaluated with the evaluation metrics. The generated synthetic dataset is divided into two parts such as, Training dataset (It is used for building the pattern tree model and test dataset (It is used for testing the web recommendation approach).

b) Performance of the W-Prefixspan Algorithm

Here, the performance of the W-prefixSpan algorithm is analyzed with the execution time, memory usage and patterns mined. At first, the training data is given to the prefixSpan and W-prefixSpan algorithm to mine the sequential patterns and then, the mining performance of the algorithms are analyzed. The values obtained from the experimentation are plotted as graphs, shown in figure 2, 3 and figure 4.

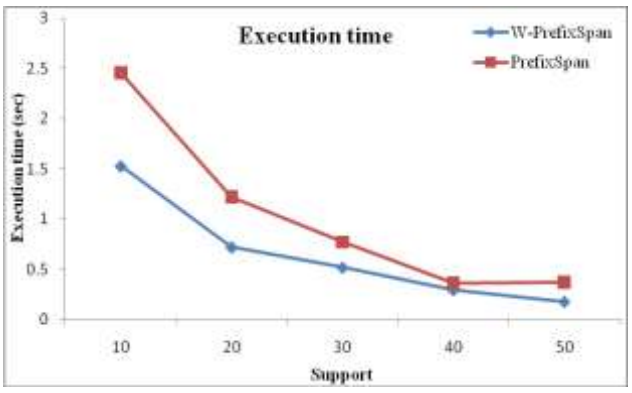


Fig. 2 : Computation time

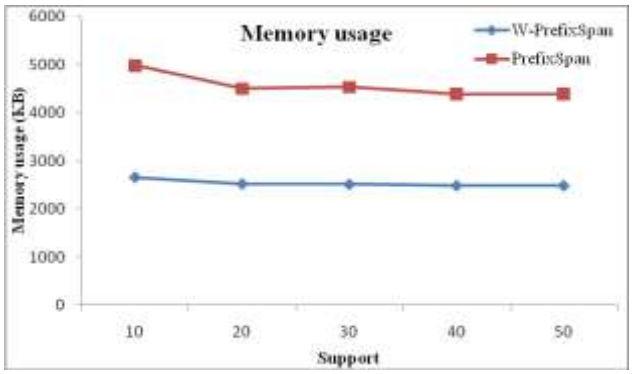


Fig.3: Memory usage

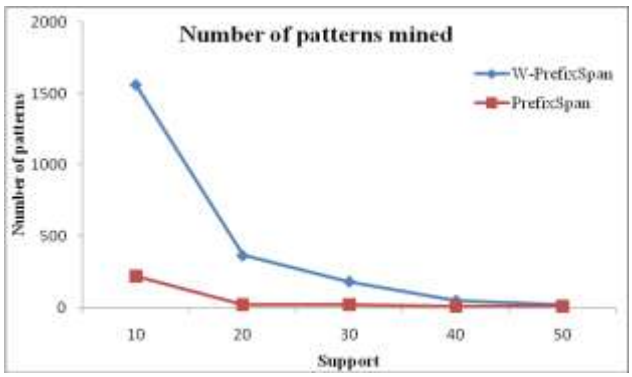


Fig.4 : Number of patterns mined

The figure 2 shows that the execution time of the W-prefixSpan algorithm is 1.5 Sec that is less compared with the time taken of the PrefixSpan algorithm. In memory usage, the W-prefixSpan algorithm needed only less than 50% memory compared with Prefixspan algorithm. Figure 4 shows that number of patterns mined using W-Prefixspan is high compared to Prefixspan algorithm.

V. Conclusion

We have proposed a web page recommendation algorithm using weighted sequential patterns and markov model. Here, we have presented W-PrefixSpan algorithm that is developed by incorporating the weightage constraints such as, spending time and recent visiting with the prefixspan algorithm. The mined weighted sequential patterns are then utilized to construct the recommendation model using the patricia trie-based tree structure. At last, markov model-based recommendation is carried out for the current users by matching the visiting path with the tree and markov model. The experimentation is done with the help of synthetic dataset and the performance of W-Prefixspan algorithm as well as web page recommendation algorithm is analyzed. From the results, the memory required for the W-prefixSpan algorithm is less than 50% of memory needed for PrefixSpan algorithm.

References Références Referencias

1. Baoyao Zhou, Siu Cheung Hui, Kuiyu Chang, "An Intelligent Recommender System using Sequential Web Access Patterns", in proceedings of IEEE Conference on Cybernetics and Intelligent Systems, 2004.
2. Daniel Mican, Nicolae Tomai, "Association-Rules-Based Recommender System for Personalization in Adaptive Web-Based Applications", Proceeding ICWE'10 Proceedings of the 10th international conference on Current trends in web engineering, pp.85-90, 2003.
3. Forsati, R., Meybodi, M.R., Rahbar, A. "An efficient algorithm for web recommendation systems", IEEE/ACS International Conference on Computer Systems and Applications, pp.579-586, 2009.
4. Faten Khalil, Jiuyong Li, Hua Wang, "Integrating Recommendation Models for Improved Web Page Prediction Accuracy", in Proceedings of the thirty-first Australasian conference on Computer science, Vol. 74, 2008.
5. Hiroshi Ishikawa, Manabu Ohta, Shohei Yokoyama, Junya Nakayama, Kaoru Katayama, "On the Effectiveness of Web Usage Mining for Page Recommendation and Restructuring", Proceeding Revised Papers from the NODe 2002 Web and Database-Related Workshops on Web, Web-Services, and Database Systems, pp.253-267, 2003.
6. Jatin D Parmar and Sanjay Garg, "Modified Web Access Pattern (mWAP) Approach for Sequential Pattern Mining", Journal of computer Science, Vol. 6, No.2, pp.46-54, June 2007.
7. Jian Pei, Jiawei Han, Mortazavi-Asl B., Pinto H., Qiming Chen, Dayal U., Mei-Chun Hsu, "PrefixSpan,: mining sequential patterns efficiently by prefix-projected pattern growth", in Proceedings of 17th International Conference on Data Engineering, 2001.
8. Qiankun Zhao and Sourav S. Bhowmick, "Sequential Pattern Mining: A Survey", Technical Report, CAIS, Nanyang Technological University, Singapore, No.118, 2003.
9. Sizu Hou, Xianfei Zhang, "Alarms Association Rules Based on Sequential Pattern Mining Algorithm", In proceedings of the Fifth International Conference on Fuzzy Systems and Knowledge Discovery, vol. 2, pp.556-560, Shandong, 2008.
10. Salvatore Orlando, Raffaele Perego and Claudio Silvestri, "A new algorithm for gap constrained sequence mining", In Proceedings of the ACM Symposium on Applied Computing, pp. 540 – 547, Nicosia, Cyprus, 2004.
11. C.P. Sumathi,R. Padmaja Valli,T. Santhanam, "Automatic Recommendation of Web Pages in Web

Usage Mining", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 09,pp.3046-3052, 2010.
12. Sarabjot Singh Anand, Bamshad Mobasher, "Intelligent Techniques for Web Personalization", ITWP 2003, LNAI 3169, pp. 1-36, Springer-Verlag Berlin Heidelberg, 2005.
13. Utpala Niranjan, R. B. V. Subramanyam and V. Khanaa, "Developing a Web Recommendation System Based on Closed Sequential Patterns", Communications in Computer and Information Science, Vol. 101, no. 1, pp. 171-179, 2010.
14. Utpala Niranjan, R.B.V. Subramanyam, V.Khana, "An Efficient System Based On Closed Sequential Patterns for Web Recommendations", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 3, No 4, May 2010.
15. Yicen Liu, Mingrong Liu, Xing Chen, Liang Xiang, Qing Yang "Automatic Tag Recommendation for Weblogs", International Conference on Information Technology and Computer Science, pp.546-549, 2009.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Layered and Feature Based Image Segmentation Using Vector Filtering

By Dr. S.Kannan & Mr. M.Sivakumar
Annamalai University, Chidambaram, Tamilnadu, India

Abstract - A Sensor is a device that reads the attribute and changes it into a signal that can be simply examined by an observer or instrument. Sensors are worked in daily objects like touch-sensitive elevator buttons, road traffic monitoring system and so on. Each sensor would carry distinctive capabilities to utilize. The objects obtained in the sensor are tracked by many techniques which have been presented earlier. The techniques which make use of the information from diverse sensors normally termed as data fusion. The previous work defined the object tracking using Multi-Phase Joint Segmentation-Registration (MP JSR) technique for layered images. The downside of the previous work is that the MP JSR technique cannot be applied to the natural objects and the segmentation of the object is also being an inefficient one.

To overcome the issues, here we are going to present an efficient joint motion segmentation and registration framework with integrated layer-based and feature-based motion estimation for precise data fusion in real image sequences and tracking of interested objects. Interested points are segmented with vector filtering using random samples of motion frames to derive candidate regions. The experimental evaluation is conducted with real image sequences samples to evaluate the effectiveness of data fusion using integrated layer and feature based image segmentation and registration of motion frames in terms of inter frame prediction, image layers, image clarity.

Keywords : *Layered image, Featured image, Image segmentation, Vector filtering.*

GJCST Classification: *I.4.6*

Strictly as per the compliance and regulations of:

Layered and Feature Based Image Segmentation Using Vector Filtering

Dr. S.Kannan [α] & Mr. M.Sivakumar [σ]

Abstract - A Sensor is a device that reads the attribute and changes it into a signal that can be simply examined by an observer or instrument. Sensors are worked in daily objects like touch-sensitive elevator buttons, road traffic monitoring system and so on. Each sensor would carry distinctive capabilities to utilize. The objects obtained in the sensor are tracked by many techniques which have been presented earlier. The techniques which make use of the information from diverse sensors normally termed as data fusion. The previous work defined the object tracking using Multi-Phase Joint Segmentation-Registration (MP JSR) technique for layered images. The downside of the previous work is that the MP JSR technique cannot be applied to the natural objects and the segmentation of the object is also being an inefficient one.

To overcome the issues, here we are going to present an efficient joint motion segmentation and registration framework with integrated layer-based and feature-based motion estimation for precise data fusion in real image sequences and tracking of interested objects. Interested points are segmented with vector filtering using random samples of motion frames to derive candidate regions. The experimental evaluation is conducted with real image sequences samples to evaluate the effectiveness of data fusion using integrated layer and feature based image segmentation and registration of motion frames in terms of inter frame prediction, image layers, image clarity.

Keywords : Layered image, Featured image, Image segmentation, Vector filtering.

I. Introduction

Layered images defined as the images taken by sensors at diverse perspectives of a general scene. Featured images are images which represent the association of the images such as points, edges, lines and contours. To identify the particular object in both types of image, many techniques have been presented earlier. The techniques which make use of the information from the sensors are termed as data fusion. Data fusion includes a well-established classification of "fusion levels" that sets diverse iterative processes of differing maturity levels. The bottom level, i.e., 0-level of "data alignment", is the registration, preprocessing, and geo-registration of images, which organizes the data for other mixture levels. Image registration, which discovers the association or the transformation between two images, thus adds to this stage in the data fusion hierarchical construction.

Author α : Professor, Annamalai University, Chidambaram, Tamilnadu, India.
Author σ : ResearchScholor, Annamalai University, Chidambaram, Tamilnadu, India. E-Mail: Kannan,Sivakum.Slm13@gmail.com

Normally, segmentation and registration are closely tangled. The main job of registration can be enhanced if some features are precisely chosen on the two images. These features might be edges, corners, and contours of the image. Registration is done by the association of feature sets which has been attained. Segmentation and registration may also be joined together in a variant structure using an active contour method. Noise removal is one of the main applications, while noise can begin from various sources and is present in nearly any image processing system. Two essential types of noises are noise created during image formation (i.e., sensor noise) and noise created during transmission (i.e., channel noise). Vector Filtering is a technique which has been used to remove the noise recognition in the digital colored images.

The main contribution of this paper is to address a real and practical problem in layered and featured images in data fusion, specifically, register and segment from different perspectives at mid-range. We propose a vector filtering method to segment and register layered images and featured images together, which, as opposed to traditional layered images, require more than one transformation for registration.

II. Literature Review

Several researches have been presented in the image segmentation process. The image segmentation is done with several techniques that have been presented earlier. Images which are mainly classified as layered image and featured image. The segmentation techniques developed individually, that is, for layered image and the featured image alone. Ping-Feng Chen, Hamid Krim ET. AL 2010 developed a technique for image segmentation for layered image using multi-phase joint segmentation technique. The downside of the project is that this technique cannot be applied to natural images.

L. M. G. Fonseca, D. Fedorov, et. Al., 2006 developed a technique for image segmentation remotely using Automatic registration and mosaicking system which allowed the image segmentation process to be done automatically. S. Jwa, U. Ozguner 2008 presented an Information-theoretic data registration for segmentation process. To implement a multi-phase process for image segmentation, H. Li and X. Tai 2007 developed a technique using level set methods to obtain a clear image quality.

Mohamed Ben Salah et. Al., 2011 used Parametric Kernel Graph Cuts method for image segmentation. Hernâni Gonçalves, José Alberto Gonçalves et. Al., 2011 developed a technique HAIRIS for image registration process using Histogram-Based Image Segmentation. Salah, M. et. Al., 2010 used an active curve objective functional with two terms: an original term and a classic length regularization term for image segmentation.

Sandberg, B, Sung Ha Kang et. Al., 2010 proposed an algorithm that automatically selects a phases to segments the image. Ghosh, P et. Al., 2010 introduced a robust image segmentation method based on a variational formulation using edge flow vectors. Mignotte, M 2010 presented a novel segmentation approach based on a Markov random field (MRF) fusion model to obtain segmentation results with simpler clustering models to achieve a more reliable and accurate segmentation result. But id does apply for artificial images. To overcome the issies, we present a vector filtering techniques to both layered and featured image segmentation and can be applied to natural images also.

III. Image Segmentation Using Vector Filtering

An image segmentation for both layered and featured images are done with vector filtering technique. vector filtering techniques treat the color image as a vector field. For a given input image, filtering format function on the idea that an image can be partitioned into small regions, each of which can be treated as stationary. Small image sections are determined using the supporting window W = {a1 , a2, . . . , aN} with the pixels ai centered around a(N+1)/2. Operating at the pixel level, spatial filtering operators replace a(N+1)/2 with the output pixel a(N+1)/2 = f(a1 , a2, . . . , aN). The architecture diagram of vector filtering technique for both layered and featured imges is shown in fig 3.1.

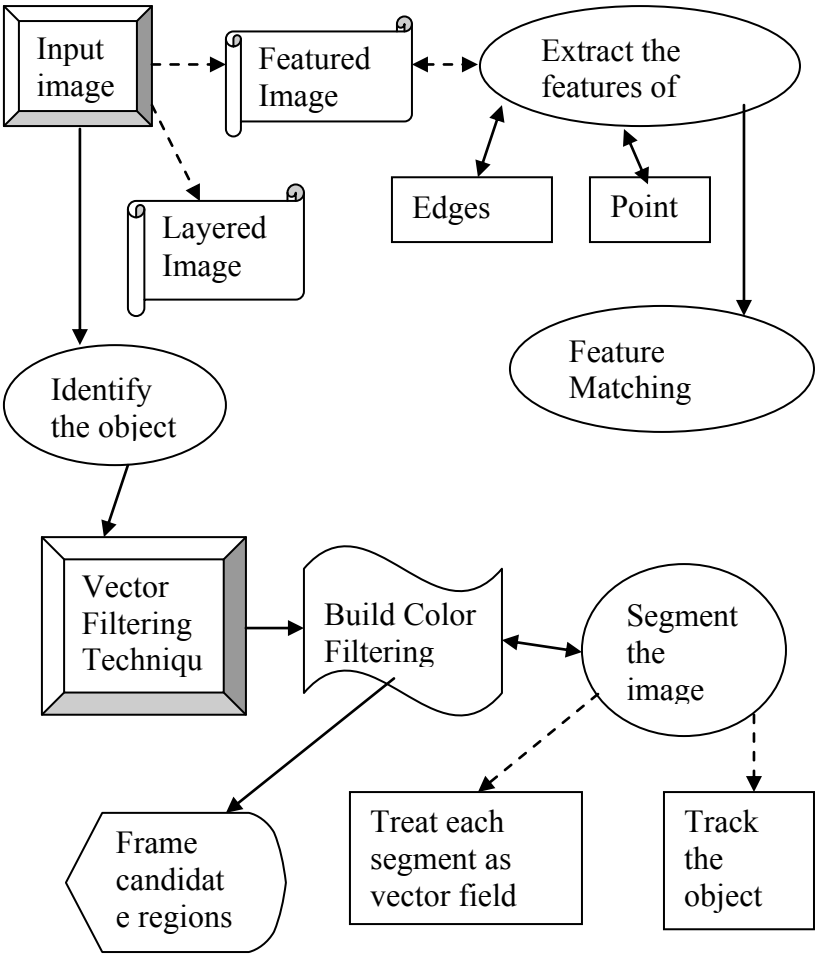


Fig 3.1: Architecture Diagram of Image segmentation using Vector Filtering Technique

For a given type of image, if it a featured image, then extract the features of the images like edges, points, and contours. Identify the interested points on the image and the match the feature with the other image to find the association between the images. If it is layered image, then identify the object which is to be tracked. Use the vector filtering technique for segmenting the image as a vector field. Then track the object till it proves its clarity. Find the motion frames of the image and find the candidate regions. Finally we get the segmented image without noise.

a) Pseudo Code of Vector Filtering Process for Image Segmentation

For a better image segmentation, here we used vector filtering technique which filters the noise efficiently from the image. The objects can also be traced based on stationary or non stationary movements. If the object is non stationary, consider the segmented region as stationary and trace the object. If the object moves out of the view, then move the window according to the position of the non stationary objects based on the features of image.

i. *Extract Features Oflayered/Featured Image*

If the input image is layered image or featured image, then it is necessary to extract the features of the images like edges, points, contours etc. Match the features of the image with other image to define the characteristics of the image and select the object in the image to be viewed.

Layered Image be L and featured image be F

Step 1 Input: Layered Image / Featured Image
Step 2 If L/F image do
Step 2.1 Extract the features like edges, points, etc.,
Step 2.2 Match the features with the image
Step 2.3 Identify the object O to be viewed
Step 3 End If

ii. *Vector Filtering Process*

The output image which we get from 3.1.1, is used here as an input for segmentation. After segmentation, treat each segmented image as a vector field and obtain the object to be traced. The output of this process is the object to be viewed by the user and segmented image with good quality.

STEP 1 Apply Vector Filtering process for the given input image after obtained features
Step 1.1 Segment the image L or F as S1, S2,….,Sn.
Step 2 For Each Si do {where i=1,2,….n}
Step 2.1 Compute the vector field v1, v2…

Step 2.2 Locate the proper size windows on images L or F
Step 2.3 Denote the windows as w
Step 2.4 Within w,
Step 2.4.1 Identify the object surface
Step 2.5 If Object O moves out of w,
Step 2.5.1 Move the window according to the feature
Step 2.6 End If
Step 3 End For
Step 4 End

Outputs: segmented Image, Object traced Segment the given input images and evaluate the vector field for each segmented image with a proper window size. Using window size, the object surface is identified and moves the window according to the movement of the segmented image of the given input image. So that it is simple to detect the traceable object

IV. Experimental evaluation

The Vector Filtering technique for image segmentation is implemented using Java. The experiments were run on an Intel P-IV machine with 2 GB memory and 3 GHz dual processor CPU. The segmentation results on 100 * 100 pixel windows of a 512 * 512 pixel image. The advantage of using a vector filtering technique is that it has an ability to handle both stationary and non stationary objects in the image and the regions inside and outside the images are automatically described.

The image layer transparency defines the different elements of an image in which imaging effects or images are applied and placed. After the image has been segmented using vector filtering technique, the transparency of the image is good when compared to an existing MP JSR technique. After segmentation is done with vector filtering, the inter frame size of the image is good in transparency by removing the noise.

The metrics used in the work involves Mean Square Error (MSE) and Peak Signal to Noise Ration (PSNR) for the image sequence without noise. The MSE is the squared error that lies between the original and segmented image. The PSNR value is the measure of peak error.

$$MSE = \sum\sum [I (a,b) – I' (a,b)] \quad \text{(eqn 4.1)}$$

Where a = 1,2,…m and b = 1,2,….n m = 512 and n = 512. From eqn 4.1 I (a,b) represents the original image whereas I' (a,b) represents the segmented image. The dimensional values are denoted by using m and n.

$$PSNR = 10 \log 10 [R^2 / MSE] \quad \text{(eqn 4.2)}$$

PSNR value is calculated in terms of decibels. The value lies between original and segmented images. PSNR ratio is used as quality of measurement Higher the PSNR rate better will be the quality of segmented image. R value is expressed in terms of 256 or 512 pixels.

V. Results and discussion

From the proposed Vector Filtering technique for both layered and featured image segmentation, it efficiently frames the candidate regions from the segmented image using vector filtering technique. By using eqn 4.1 and 4.2, the PSNR values for the segmented image are computed to evaluate the performance and the quality of the segmented image. Table 5.1 describes the image quality after it get segmented using the proposed Vector Filtering (VF) technique for both layered and featured image segmentation and an existing Multiphase Joint Segmentation-Registration and Object Tracking for Layered Images.

Image Size	PSNR values (db)	
	Proposed VF image segmentation	Existing Multi-phase Joint Image segmentation
512 *512	31.45	45.53

Table 5.1: PSNR comparison table

The comparison for image segmentation by the proposed Vector Filtering technique for both layered and featured image segmentation and an existing Multiphase Joint Segmentation-Registration and Object Tracking for Layered Images is explained in the comparison graphs.

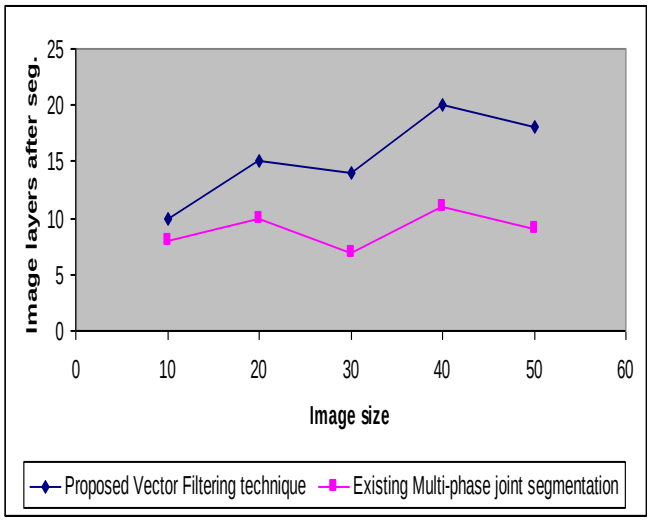


Fig 5.1: Image size vs Image Layers

Fig 5.1 describes the layers of the given image after segmentation process. Even the image size increases in pixels, the layered form of image after segmentation is good in the proposed Vector Filtering technique for both layered and featured image segmentation. Depends on the image size, the layered image has been formed well in the proposed Vector Filtering technique for both layered and featured image segmentation, since the variance would be 20-25% high.

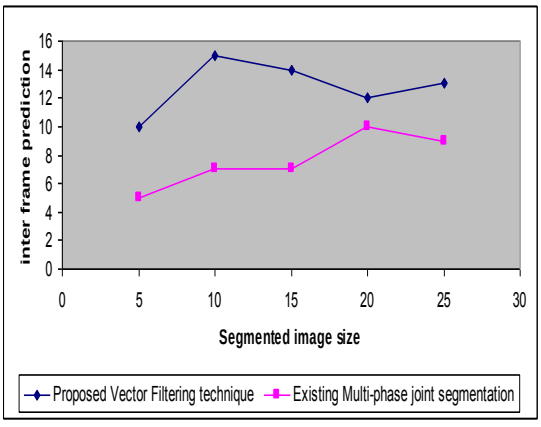


Fig 5.2 : segmented image size vs Inter frame prediction

Fig 5.2 described the inter frame prediction of segmented image. The inter frame prediction is used to predict the segmented image quality to evaluate whether it achieves high compression rate or not. In the proposed Vector Filtering technique for both layered and featured image segmentation, the inter frame prediction is good. So, the segmented image quality is also being good compared to an existing Multiphase Joint Segmentation-Registration and Object Tracking for Layered Images.

Finally, observed that the experimental results and the table 5.1 shows that the image segmentation process for both layered and featured image performed better in terms of image layer, image quality with less noise level to obtain a better image results. The quality of the image is 20-25% high in the proposed Vector Filtering technique for both layered and featured image segmentation.

VI. Conclusion

The Vector Filtering technique for both layered and featured image segmentation proposed in this paper efficiently done the segmentation process in both types of image without any noise. The vector filtering technique used for image segmentation outperforms better in terms of segmented image transparency, image quality. Since the vector filtering technique treated each segmented image as vector field, the proposed Vector Filtering technique for both layered and featured image segmentation successfully had done the image segmentation and registration process. The user can easily get the object in a clear manner by tracking the object in the segmented image. The segmented image quality is 15-20% high compared to an existing Multiphase Joint Segmentation-Registration and Object Tracking for Layered Images. The experimental results have shown that the proposed Vector Filtering technique for both layered and featured image segmentation achieved better results when compared to an existing Multiphase Joint Segmentation-Registration and Object Tracking for Layered Images.

References références referencias

1. Ping-Feng Chen, Hamid Krim ET. AL., "Multiphase Joint Segmentation-Registration and Object Tracking for Layered Images", IEEE transactions on image processing, vol. 19, no. 7, july 2010.
2. L. M. G. Fonseca, D. Fedorov, B. S. Manjunath, C. Kenney, E. Castejon, and J. S. de Medeiros, "Automatic registration and mosaicking system for remotely sensed imagery," *Revista Brasileira de Cartografia*, vol. 58, pp. 49–61, 2006.
3. S. Jwa, U. Ozguner, and Z. Tang, "Information-theoretic data registration for uav-based sensing," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 5–15, Mar. 2008.

4. H. Li and X. Tai, "Piecewise constant level set methods for multiphase motion," *Int. J. Numer. Anal. Modeling*, vol. 4, pp. 274–293, 2007.
5. Mohamed Ben Salah, *Member, IEEE*, Amar Mitiche et. Al., "Multiregion Image Segmentation by Parametric Kernel Graph Cuts", IEEE transactions on image processing, vol. 20, no. 2, february 2011.
6. Hernâni Gonçalves, José Alberto Gonçalves, and Luís Corte-Real, "HAIRIS: A Method for Automatic Image Registration Through Histogram-Based Image Segmentation", IEEE transactions on image processing, vol. 20, no. 3, march 2011.
7. Jianping Fan, Yi Shen, Chunlei Yang, and Ning Zhou, "Structured Max-Margin Learning for Inter-Related Classifier Training and Multilabel Image Annotation", IEEE transactions on image processing, vol. 20, no. 3, march 2011.
8. G. Unal, H. Krim, and A. Yezzi, "Fast incorporation of optical flow into active polygons," *IEEE Trans. Image Process.*, vol. 14, no. 6, pp. 745–759, Jun. 2005.
9. D. Cremers and S. Soatto, "Motion competition: A variational approach to piecewise parametric motion segmentation," *Int. J. Comput.Vis.*, vol. 62, no. 3, pp. 249–265, 2005.
10. G. Unal, A.Yezzi, and H. Krim, "Information-theoretic active polygons for unsupervised texture segmentation," *Int. J. Comput. Vis.*, vol. 62, pp. 199–220, 2005.
11. Mignotte, M, "A Label Field Fusion Bayesian Model and Its Penalized Maximum Rand Estimator for Image Segmentation", Image processing,IEEE transaction on june 2010 .
12. Ghosh, P.; Bertelli, L.; Sumengen, B.; Manjunath, B.S, "A Nonconservative Flow Field for Robust Variational Image Segmentation", Image processing,IEEE transaction on feb 2010 .
13. Sandberg, B.; Sung Ha Kang; Chan, T.F., "Unsupervised Multiphase Segmentation: A Phase Balancing Model", Image processing,IEEE transaction on jan 2010 .
14. Ben Salah, M.; Mitiche, A.; Ben Ayed, I, "Effective Level Set Image Segmentation With a Kernel Induced Data Term", Image processing,IEEE transaction on jan 2010 .

Global Journal of Computer Science and Technology Volume XII Issue IX Version I April 2012
14

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Robust Digital Watermarking for Color Images Using Fuzzy Vault

By Priyanka D. Godase, Snehal B. Kale, Sonika S. Shelke, S.M.Sangve & S.P.Deshmukh

Dnyanganga College of Engineering and Research; Pune, Maharashtra, India

Abstract - The advent of the Internet has resulted in many new opportunities for the creation and delivery of content in digital form. Applications include electronic advertising, real-time video and audio delivery, digital repositories and libraries, and Web publishing. An important issue that arises in these applications is the protection of the rights of all participants. It has been recognized for quite some time that current copyright laws are inadequate for dealing with digital data. This has led to an interest towards developing new copy deterrence and protection mechanisms. One such effort that has been attracting increasing interest is based on digital watermarking techniques. Digital watermarking is the process of embedding information into digital multimedia content such that the information (the watermark) can later be extracted or detected for a variety of purposes including copy prevention and control. Digital watermarking has become an active and important area of research, and development. Key research problem that we still face today is the development of truly robust, transparent and secure watermarking technique for different digital media including images, video and audio. So we propose a new robust digital watermarking system based on DCT and fuzzy vault techniques.

Keywords : *DCT; color components; color image; fuzzy vault; Digital watermarking.*

GJCST Classification: *I.2.3*

Strictly as per the compliance and regulations of:

Robust Digital Watermarking for Color Images Using Fuzzy Vault

Priyanka D. Godase[α], Snehal B. Kale[σ], Sonika S. Shelke[ρ], S.M.Sangve[ω] & S.P.Deshmukh[¥]

Abstract - The advent of the Internet has resulted in many new opportunities for the creation and delivery of content in digital form. Applications include electronic advertising, real-time video and audio delivery, digital repositories and libraries, and Web publishing. An important issue that arises in these applications is the protection of the rights of all participants. It has been recognized for quite some time that current copyright laws are inadequate for dealing with digital data. This has led to an interest towards developing new copy deterrence and protection mechanisms. One such effort that has been attracting increasing interest is based on digital watermarking techniques. Digital watermarking is the process of embedding information into digital multimedia content such that the information (the watermark) can later be extracted or detected for a variety of purposes including copy prevention and control. Digital watermarking has become an active and important area of research, and development. Key research problem that we still face today is the development of truly robust, transparent and secure watermarking technique for different digital media including images, video and audio. So we propose a new robust digital watermarking system based on DCT and fuzzy vault techniques.

Keywords : DCT; color components; color image; fuzzy vault; Digital watermarking.

I. Introduction

Information hiding can be mainly divided into three processes - cryptography, steganography and watermarks. Cryptography is the process of converting information to an unintelligible form so that only the authorized person with the key can decipher it. As many advances were made in the field of communication it became rather simple to decrypt a cipher text. Hence more sophisticated methods were designed to offer better security than what cryptography could offer. This led to the discovery of stenography and watermarking. Stenography is the process of hiding information over a cover object such that the hidden information cannot be perceived by the user. Thus even the existence of secret information is not known to the attacker. Watermarking is closely related to steganography, but in watermarking the hidden information is usually related to the cover object. Hence it is mainly used for copyright protection and owner authentication.

Digital watermarking is the process of embedding information into digital multimedia content such that the information (the watermark) can later be extracted or detected for a variety of purposes including copy prevention and control. Digital watermarking has become an active and important area of research, and development. Key research problem that we still face today is the development of truly robust, transparent and secure watermarking technique for different digital media including images, video and audio. So we propose a new robust digital watermarking system based on DCT and fuzzy vault techniques

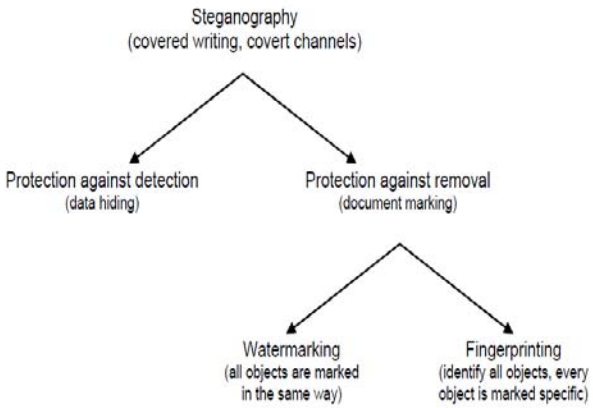


Figure 1: Types of Steganography

II. Principle of Watermarking

A watermarking system is usually divided into three distinct steps, embedding, attack and detection. In embedding, an algorithm accepts the host and the data to be embedded and produces a watermarked signal. The watermarked signal is then transmitted or stored, usually transmitted to another person. If this person makes a modification, this is called an attack. There are many possible attacks. Detection is an algorithm which is applied to the attacked signal to attempt to extract the watermark from it. If the signal was not modified during transmission, then the watermark is still present and it can be extracted. If the signal is copied, then the information is also carried in the copy. The embedding takes place by manipulating the content of the digital

Author α σ ρ : UG Student, Department of Computer, Dnyanganga College of Engineering and Research; Pune, Maharashtra, India.
Author ω : Head of the Department of Computer, Dnyanganga College of Engineering and Research; Pune, Maharashtra, India.
Author ¥ : Assistant Professor, Department of Computer, Dnyanganga College of Engineering and Research; Pune, Maharashtra, India.
E-mail α : priyanka.godase@gmail.com
E-mail σ : manusnehal@gmail.com
E-mail ρ : sonika_shelke@yahoo.com
E-mail ω : sunilsangve@gmail.com
E-mail ¥ : sayaleedeshmukh@yahoo.com

data, which means the information is not embedded in the frame around the data, it is carried with the signal itself.

III. Digital Watermarking Technology

As an Emerging Interdisciplinary application Technology, Digital Watermarking Involves the Ideas and Theories of Different Subject Coverage Such as Signal Processing, Cryptography, Probability Theory, Network Theory Algorithm Design, and other techniques.

a) Classfication of Digital Watermarking

i. *Visible*

The watermark is visible which can be a text or a logo used to identify the owner. Any text or logo to verify or hide content.

ii. Invisible

The watermark is embedded into the image in such a way that it cannot be perceived by human eye. It is used to protect the image authentication and prevent it from being copied.

Invisible watermark can be further divided into three types,

a. *Robust Watermarks*

Invisible watermark cannot be manipulated without disturbing the host signal. This is by far the most important requirement of a watermark. There are various attacks, unintentional (cropping, compression, scaling) and unintentional attacks which are aimed at destroying the watermark. So, the embedded watermark should be such that it is invariant to various such attacks. They are designed to resist any manipulations that may be encountered. All applications where security is the main issue use robust watermarks.

b. *Fragile Watermarks*

They are designed with very low robustness. They are used to check the integrity of objects.

c. *Public and Private Watermark*

They are differentiated in accordance with the secrecy requirements for the key used to embed and retrieve watermarks. If the original image is not known during the detection process then it is called a public or a blind watermark and if the original image is known it is called a non blind watermark or a private watermark.

IV. System Architecture

a) Embedding Watermark

This module is designed to insert the fuzzy vault into the host image. The fuzzy vault is combination of set of genuine points and set of chaff points. Secret data is used to construct the polynomial.

In this module first we extract minutiae features from fingerprint image. And these features are further used to project the polynomial then set of genuine points and chaff points is calculated. And the union of set of genuine points and chaff points is nothing but the fuzzy vault.For inserting the fuzzy vault into an image DCT is applied on that image and to get the watermarked image IDCT is applied.

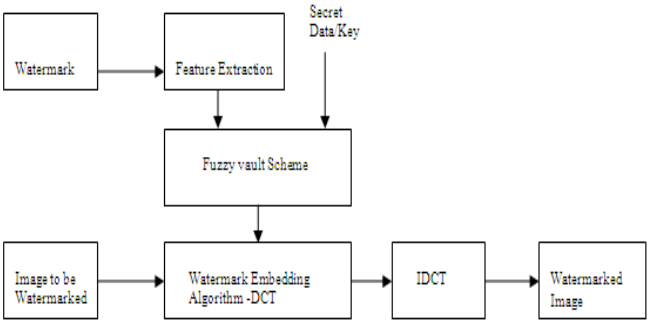


Figure 2: Embedding watermark

b) Extracting Watermark

In this module secret data is extracted by using the watermarked image and fingerprint image to validate the user. Again the DCT is applied on the watermarked image and fuzzy vault is extracted from that image. And minutiae features are extracted from the fingerprint image. By comparing the minutiae features and fuzzy vault candidate points are calculated. And then applying Langrange Interpolation CRC is calculated. If the CRC is correct then we get the secret data that proves that the user is valid user.

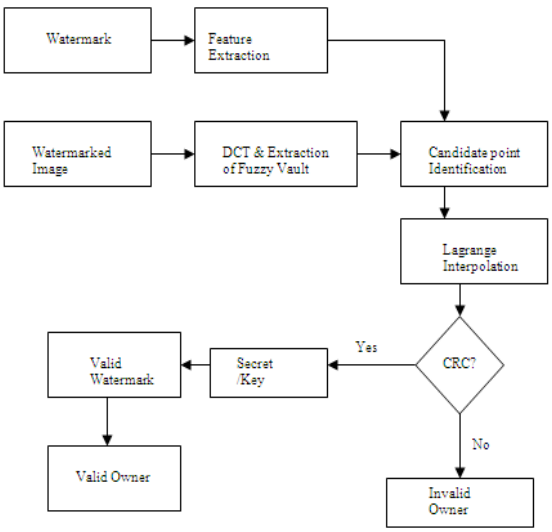


Figure 3: Extracting watermark

c) Main Algorithms

i. *Encoding Algorithm*

1. Choose the secret value or the cryptographic key S
2. Generate the cyclic redundancy check value CRC

3. Construct 144-bits data SC by concatenation of S and CRC
4. Extract k-eigenvectors E to arrive at the 16-bit locking data unit u
5. Construct a polynomial P with SC: SC can be represented as a polynomial with 9 (=144/16) coefficients for the degree D = 8.

$$P(u) = c8u\hat{}8 + c7u\hat{}7 + ...+ c1u + c0$$

6. Generate genuine points set G with ui

$$G = \{(u1, P(u1)), (u2, P(u2)),.., (uN, P(uN))\}$$

$$ui \neq uk ,i = 1,2,...,N$$

7. Generate chaff points set C

$$C = \{(c1, d1), (c2, d2), ..., (cM, dM)\}$$

$$dj \neq p(cj), j = 1, 2,\ldots M$$

8. Generate vault set VS: VS = C U G

$$VS = \{(V1, W1), (V2, W2),.., (VN+M, WN+M)\}$$

ii. *Decoding Algorithm*

1. Confirm the vault set VS

$$VS = \{(V1, WI), (V2, W2),.., (VN+M, WN+M)\}$$

2. Extract N-features u*

$$u*1, u*2, \ldots, u*N$$

3. Find K-candidate points: If any u*i is equal to Vj, the corresponding vault point is added to the list of candidate points, where K< N.
4. Find all possible combinations ofD+1 points among t he list of candidate points, resulting in C (K, D+1) combinations.
5. Reconstruct polynomial p*(u): a specific combination set given as:

$$L= \{(v1, w1), (v2, w2), (v3, w3) ,\ldots ,(vD+1,wD+1)$$

the corresponding interpolating polynomial method is the Lagrange method.

$$p^* (u) = C*8u\hat{}8+C^* 7u\hat{}7 +...+C*1u+C*0$$

6. Check CRC
7. Decode secret value S

iii. *Steps for Getting Watermarked Image*

1. Divide host image in to 8x8 blocks.
2. Calculate DCT for each block from left to right and top to bottom.
3. Compress image using Quantization technique.(Quantization: It is the Process of approximation.)
4. By using IDCT we get original Watermarked image.

V. Application

In this project, an efficient blind digital image watermarking algorithm using mapping technique is presented. The algorithm can embed or hide an entire image or pattern (logo) directly into the original image. The embedding process is based on changing the selected DCT coefficients of the host image to odd or even values depending on the binary bit value of watermark DCT coefficients. The algorithm is tested for fingerprint image embedded with a face watermark. It is demonstrated that the watermarking algorithm offers a significant advantage of providing biometric image compression and authentication without introducing any significant degradation in the image quality. Moreover the watermarking scheme is blind and does not require any additional data for logo extraction.

Applications of image watermarking:

a) Copyright Protection

This is by far the most prominent application of watermarks. With tons of images being exchanged over insecure networks every day, copyright protection becomes a very important issue. Watermarking an image will prevent redistribution of copyrighted images.

b) Authentication

Sometimes the ownership of the contents has to be verified. This can be done by embedding a watermark and providing the owner with a private key which gives him an access to the message. ID cards, ATM cards, credit cards are all examples of documents which require authentication.

c) Broadcast Monitoring

As the name suggests broadcast monitoring is used to verify the programs broadcasted on TV or radio. It especially helps the advertising companies to see if their advertisements appeared for the right duration or not.

d) Content Labeling

Watermarks can be used to give more information about the cover object. This process is named content labeling.

e) Tamper Detection

Fragile watermarks can be used to detect tampering in an image. If the fragile watermark is degraded in any way then we can say that the image or document in question has been tampered.

f) Digital Fingerprinting

This is a process used to detect the owner of the content. Every fingerprint will be unique to the owner.

g) Content Protection

In this process the content stamped with a visible watermark that is very difficult to remove so that it can be publicly and freely distributed.

VI. Conclusion

With the popularity of the network, the safety communication issue of digital product becomes an important and urgent research topic. The basic principles and algorithms of the digital watermarking technology are discussed, and the DCT algorithm is selected to do the application test of digital image copyright protection.

The experiment proves that DCT-based watermark can well withstand a variety of image processing, and the watermark can survive after compression, cropping, and other attacks.

Digital watermarking technology can provide a new way to protect the copyright of multimedia information and to ensure the safe use of multimedia information. Comparing to the traditional information security technology, digital watermarking technology has its own advantages in the multimedia information security protection. Then it can meet the application need in many aspects and has a bright development prospect.

References Références Referencias

1. David Salomon. Data Compression - The Complete Reference 3rd Edition, Springer, 2004
2. N. Ahmed, T. Natarajan, and K. R. Rao, Discrete cosine transform, IEEE Trans. Comput., vol. C-23, Jan. 1974, 90-93.
3. G. A. Juels and M. Sudan, "A Fuzzy Vault Scheme," *Proceedings of IEEE International Symposium on Information Theory*, 2002.
4. B. D. Tseng and W. C. Miller. On computing the discrete cosine transform, IEEE Trans. Computing, vol.C-27: July 1976, 966-968.
5. U. Uludag and A. K. Jain, "Securing fingerprint template: Fuzzy vault with helper data," in IEEE Workshop on Privacy Research In Vision, June 2006, p. 163.
6. X. Chen, J. Tian, and X. Yang, "A novel algorithm for distorted fingerprint matching based on fuzzy features match," in AVBPA 2005. 2005, vol. LNCS 3546, pp. 665-673, Springer-Verlag.
7. D.P. Mital and E.K. Teoh, "An automated matching technique for fingerprint identification," in KES 1997, May 1997, vol. 1,pp. 142-147.
8. S. Na K.D. Yu and T.Y Choi, "A fingerprint matching algorithm based on radial structure and a structure-rewarding scoring strategy," in AVBPA 2005. 2005, vol. LNCS 3546, pp. 656-664, Springer-Verlag.
9. Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in Eurocrypt 2004, 2004.
10. J. Jeffers and A. Arakala, "Minutiae-based structures for a fuzzy vault," in 2006 Biometrics Symposium, 2006.

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Computer-Based Decision Support System: A Study of Akanu Ibiam Federal Polytechnic

By Ezeorah Ezekiel U, Robert Ben Joshua, Eguzo Chimezie V & Ejighikemenwa Nwevo

Akanu Ibiam Federal Polytechnic, Unwana

Abstract - This work is aimed at developing a decision support system to improve the decision-making capacity of administrators of our case study (Akanu Ibiam Federal Polytechnic, Unwana) and other parties. The system will help them to develop their administrative skill in decision making and resource management programs. The development framework is divided into three levels that employed Web-based application at the data collection level for collection of student statistics which is the primary data. Database application was used at the Processing level to provide administrative utility for data flow control and storage while the output level depends on a spreadsheet application for summary and advisory purpose. The research though not yet completely explored, its benefits will aid in the general management of the school system.

Keywords : *Database, Statistics, Students, Web, Spreadsheet, Application.*

GJCST Classification: *H.2,H.2.4,H.4.1*

Strictly as per the compliance and regulations of:

Computer-Based Decision Support System: A Study of Akanu Ibiam Federal Polytechnic

Ezeorah Ezekiel U [α], Robert Ben Joshua [σ], Eguzo Chimezie V [ρ] & Ejighikemenwa Nwevo [ω]

Abstract - This work is aimed at developing a decision support system to improve the decision-making capacity of administrators of our case study (Akanu Ibiam Federal Polytechnic, Unwana) and other parties. The system will help them to develop their administrative skill in decision making and resource management programs. The development framework is divided into three levels that employed Web-based application at the data collection level for collection of student statistics which is the primary data. Database application was used at the Processing level to provide administrative utility for data flow control and storage while the output level depends on a spreadsheet application for summary and advisory purpose. The research though not yet completely explored, its benefits will aid in the general management of the school system.

Keywords: Database, Statistics, Students, Web, Spreadsheet, Application.

I. Introduction

According to Wikipedia, *"A decision support system (DSS) is a computer-based information system that supports business or organizational decision-making activities. DSSs serve the management, operations, and planning levels of an organization and help to make decisions, which may be rapidly changing and not easily specified in advance. DSSs include knowledge-based systems. A properly designed DSS is an interactive software-based system intended to help decision makers compile useful information from a combination of raw data, documents, personal knowledge, or business models to identify and solve problems and make decisions".[1]* Computerized decision support systems became practical with the development of minicomputers, timeshare operating systems and distributed computing [2].

Akanu Ibiam Federal Polytechnic, is a tertiary institution in Nigeria controlled by National Board for Technical Education for awarding of National Diplomas and Higher National Diplomas on various courses offered by the institution. The research framework presented in this paper is aimed at serving as a decision support system for the institution's management decision, the application is not limited to the management team alone, it is also designed to transcend to other decision areas like, the student affairs unit, Heads of Department, Deans of Schools and others inclusive. The decision of the institution's management team is completely controlled by the information gathered from different variables which include student statistics, monetary inventory and many more. This is the first research on this dimension for this case study and still being explored, so our concentration is currently on developing a decision support framework from student statistics gathered through a web based portal system, with a combination of some query and database software. The system is a combination of different level of applications that are interconnected for proper functionality. Fig 1.0 is a model approach followed for the development of our decision support system.

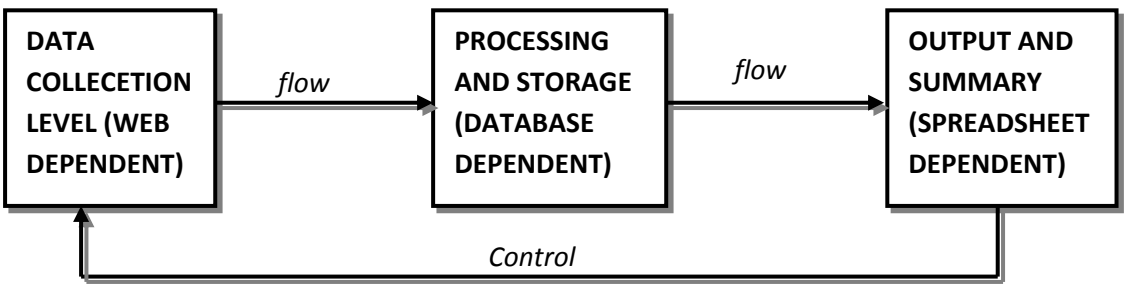


Fig 1.0 : Decision Support System Model

Author α : Computer Science Department.
Author σ ρ : Electrical/Electronic Engineering Department.
Author ω : Project Manager IT Center Akanu Ibiam Federal Polytechnic, Unwana.
E-mail α : Euezeorah@yahoo.com,
E-mail σ : bjrobert@akanuibiampoly.edu.ng
E-mail ρ : chimaxcorporations@gmail.com
E-mail ω : ejnwevo@gmail.com

II. Development Framework

The system is not entirely a software program, but a combination of different software, hardware and humanware levels. We divided the framework into three levels: Data collection, Processing and Summary levels.

From fig 1.0, the framework functions a directional architecture where the last level forms a control for the program.

DATA COLLECTION (Web-based Level): The Internet era has taken information-sharing to new heights, allowing billions of users to share information on the World Wide Web [3]. The World-wide Web and global Internet provided a technology platform for further extending the capabilities and deployment of computerized decision support. This system uses tools like Internet Explorer, Mozilla Firefox, and Netscape accessible to analyst for data management. The clients and servers must certify all the condition of network connectivity like the TCP/IP protocols [3] Considering the potentials of the World Wide Web, this category was developed to use the institution's web portal as a data cache for gathering students' statistics. The web portal is designed for management of student fee payments and registration, but for the purpose of this research work, we used this portal to form a common synergy between the web application and a database program for information collection. On the completion of the forms presented by the web portal application during student's registration, the database program queries these data according to the required fields. Fig 2.0 shows the school web application used for student information collection.

PROCESSING AND STORAGE (Management Access Level): At this level, management tools are employed for extraction of information according the required properties. This involves the use of spreadsheet and database programs for storage purposes. The method of data collection, informed our choice of web based system were processing and data analysis can be grouped in field of different properties. According to Power in www.dssresources.com, a data-driven DSS emphasizes access to and manipulation of a time-series of internal company data and sometimes external and real-time data [3]. Such data collected from our web portal is accessed using some simple file systems controlled by query and retrieval tools to provide the most elementary level of functionality. The data storage system is designed to allow the manipulation of data by computerized conditions tailored to a specific task and setting. The system outputs a summary when certain conditions are met, exceeded or approached. Fig 3.0 shows a conditional query of our database to output some required properties in a spreadsheet format using Microsoft Excel application.

OUTPUT AND SUMMARY (DECISION LEVEL): This level provides a summary of the outcomes from data collection and processing. In this category, every required condition provides details of its output fields. For instance, fig 3.0 is a detailed summary of the students' statistics which is an efficient support system for decision making purposes extracted from the processing and storage level. Spreadsheet application is employed at this stage to enable a detailed extraction of the needed information. Keys were used to help in the understanding of some coded formats because the summary is completely meant to be accessed by the decision makers who may not be computer analysts but understand basic statistical distribution of data.

III. Applications

This system finds its application in almost all the units of the institution. For instance, the academic unit can use the system for admission, accreditation and population control purposes. The administrative unit can employ this system for staff, equipment and distribution purposes. The management unit can use this system to develop a more informed relationship with the control organ being the NBTE. Many researchers have developed similar system for several applications; some have presented result for land use and agricultural purposes, Canadian National Railway System, etc[1]

IV. Benefits

The research is aimed at equipping decision markers with an accurate, in time and up-to-date information. It will provide many benefits which include and not limited to

- Improving efficiency of the system
- Enhancing the decision making process
- Developing a more efficient organizational control
- Proving avenue to faster problem solving in the institution
- Facilitate communication
- Promote learning, teaching and training
- Reveal new approach for improvement of the Decision Support System
- Automate managerial process.

V. Conclusion

The result as outlined in fig 4.0 is currently employed for the next session admission process, although challenges that will lead to improvement of the system are being anticipated. Every arrangement has been made to cushion foreseen developments as can be projected at this stage. Hence a higher version of the DSS program for more robust data control is currently being analyzed.

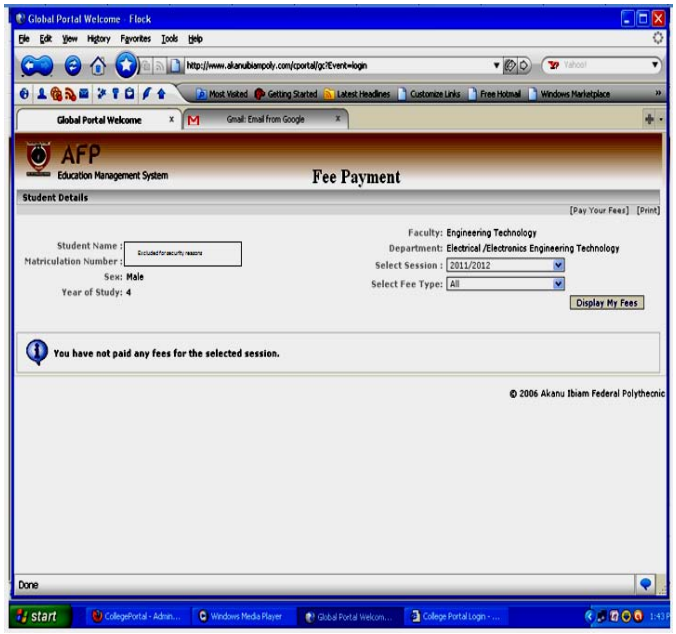


Fig 2 : Data Collection showing Fee Payment

Fig 3 : Data Processing using Management Utility

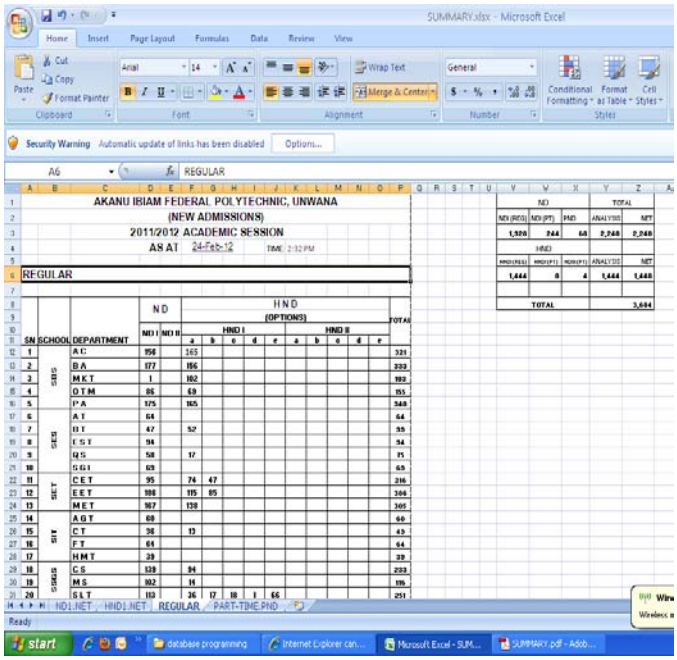


Fig 4a : Output and Summary

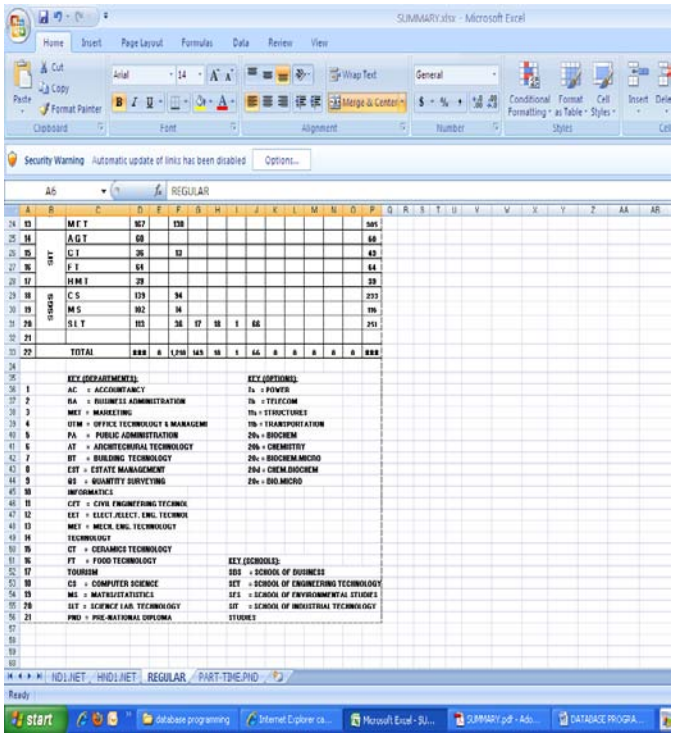


Fig 4b : Output and Summary (contd.)

References Références Referencias

1. Decision support system - Wikipedia, the free encyclopedia http://en.wikipedia.org/wiki/Decision_ support_system
2. D. J. Power: A Brief History of Decision Support Systems -http://dssresources.com/history/dsshistory.html
3. Abhijit A. Pol, Ravindra K. Ahuja-Developing Web-Enabled Decision Support Systems Using

Access, VB .NET, and ASP .NET-Dynamic Ideas Belmont, Massachusetts, WWW information and orders:http://www.dynamic-ideas.com

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Sensor Data Encryption Protocol for Wireless Network Security

By Bharat Singh, Parvinder Singh & Dr. V.S. Dhaka

Central University of Rajasthan, Ajmer, Rajasthan

Abstract - Wireless Sensor Network consisting of a large number of sensor nodes that connected through wireless media has emerged as a ground breaking technology that offers unprecedented ability to monitor the physical word accurately. The privacy preservation is an important issue in wireless sensor network. Developing effective security solutions for wireless sensor networks are not easy due to limited resources. In this paper we propose new techniques for the purpose of security in wireless sensor network called as SDEP sensor data encryption protocol. In the scheme we use the RC 6 method for the purpose of encryption and decryption. RC 6 provide best confusion and diffusion properties with the less computational overhead. In order to confirm effectiveness of SDEP, a comparative performance evaluation with AES and RC 5 algorithms are presented in terms of memory requirement and execution time criteria. Our proposed scheme provides better performance than AES and RC 5 in the term of execution time and total memory requirement. We also provide simulation results for proposed method in the term of overhead and energy according to this result SDEP is strong block cipher for wireless sensor networks.

Keywords : *SDEP, Security, RC 6 cryptography, AES.*

GJCST Classification: *C.2.0,D.4.6*

Strictly as per the compliance and regulations of:

Sensor Data Encryption Protocol for Wireless Network Security

Bharat Singh [α], Parvinder Singh [α] & Dr. V.S. Dhaka [σ]

Abstract - Wireless Sensor Network consisting of a large number of sensor nodes that connected through wireless media has emerged as a ground breaking technology that offers unprecedented ability to monitor the physical word accurately. The privacy preservation is an important issue in wireless sensor network. Developing effective security solutions for wireless sensor networks are not easy due to limited resources. In this paper we propose new techniques for the purpose of security in wireless sensor network called as SDEP sensor data encryption protocol. In the scheme we use the RC 6 method for the purpose of encryption and decryption. RC 6 provide best confusion and diffusion properties with the less computational overhead. In order to confirm effectiveness of SDEP, a comparative performance evaluation with AES and RC 5 algorithms are presented in terms of memory requirement and execution time criteria. Our proposed scheme provides better performance than AES and RC 5 in the term of execution time and total memory requirement. We also provide simulation results for proposed method in the term of overhead and energy according to this result SDEP is strong block cipher for wireless sensor networks.

Keywords : *SDEP, Security, RC 6 cryptography, AES;*

I. Introduction

A wireless Sensor Network is simple defined as a large collection of sensor nodes, equipped with its own sensor, processor and radio transceiver. A wireless sensor network has been widely used in different application areas to know the battlefield situation data, monitoring building parameters and reports about malfunction in a system.

The evolution leading to RC6 has provided a simple cipher yielding numerous evaluations and adequate security in a small package. After describing the structure of the algorithm, the prominent goal that stands out is simplicity. Through this simplicity, multiple evaluations have been performed, including AES-related evaluations, which will be discussed at a high level, due to complexity and number of articles. The fact that such a small, simple algorithm contended for AES with such high security requirements is noteworthy.

The main objective of our approach is to provide better performance than AES and RC 5 in the term of execution time and total memory requirement. We also provide simulation results for proposed method in the term of overhead and energy according to this result SDEP is strong block cipher for wireless sensor networks. In this new technique for the purpose of security in wireless sensor network called as SDEP sensor data encryption protocol. In the scheme we use the RC 6 method for the purpose of encryption and decryption. RC 6 provide best confusion and diffusion properties with the less computational overhead. In order to confirm effectiveness of SDEP, a comparative performance evaluation with AES and RC 5 algorithm are presented in terms of memory requirement and execution time criteria. Our proposed scheme provides better performance than AES and RC 5 in the term of execution time and total memory requirement.

II. Attacks on Wireless Sensor Network

Wormhole Attacks: In the wormhole attack an adversary builds a virtual tunnel through a low latency link that takes the messages from one part of the network and forwards them to another. The simplest case of this attack is when one node is located between two other nodes that are forwarding. However, wormhole attacks commonly involve two distant nodes that are colluded to underestimate the distance between them and forward packets through an external communication channel that is only available to the adversary.

Sinkhole Attacks: In a sinkhole attack, the adversary's goal is to lure nearly all the traffic from a particular area through a compromised node, creating a metaphorical sinkhole with the adversary at the center. Sinkhole attacks typically work by making a compromised node look especially attractive to surrounding nodes with respect to the routing algorithm. Effectively, the adversary creates a large "sphere of influence", attracting all traffic destined for a base stations from nodes several hops away from the compromised node.

Subversion of a Node: A particular sensor might be captured, and information stored on it (such as the key) might be obtained by an adversary. If a node has been compromised then how to exclude that node, and that node only, from the sensor network is at issue defines an efficient way to do so.

Physical Attacks: Sensor networks often operate in hostile environments. In those environments, the size of the nodes plus the unattended operation mode contributes to make them very vulnerable to physical

Author α : M.Tech. Scholar Central University of Rajasthan, Ajmer, Rajasthan.
Author σ : Prof. & Dean Rajasthan Institute of Engineering &Technology, Jaipur Rajasthan.

attacks. In contrast to other types of attacks, physical attacks destroy the nodes permanently, thus, their loss is irreversible. For instance, an adversary could extract cryptographic keys, alter the node's circuitry, and reprogram it or replace it with malicious nodes.

Passive Information Gathering: An intruder with an appropriately powerful receiver and well designed antenna can easily pick off the data stream. Interception of the messages containing the physical locations of sensor nodes allows an attacker to locate the nodes and destroy them. Besides the locations of sensor nodes, an adversary can observe the application specific content of messages including message IDs, timestamps and other fields. To minimize the threats of passive information gathering, strong encryption techniques needs to be used.

False Node and Malicious Data: An intruder might add a node to the system that feeds false data or prevents the passage of true data. Such messages also consume the scarce energy resources of the nodes. This type of attack is called "sleep deprivation torture".

The Sybil Attack: In a Sybil Attack, a single node presents multiple identities to other nodes in the network. They pose a significant threat to geographic routing protocols, where location aware routing requires nodes to exchange coordinate information with their neighbors to efficiently route geographically addressed packets. Authentication and encryption techniques can prevent an outsider to launch a Sybil Attack on the sensor network.

Acknowledgement spoofing: Some routing algorithms require the use of acknowledgement signals (ACK). In this case, an adversary could spoof this signal in response to the packets that the adversary listens to. This results in convincing the transmitting node that a weak link is strong. Thus, an adversary could perform a selective forwarding attack after spoofing ACK signals to the node that the adversary intends to attack. Attacks to Data Aggregation Techniques Data aggregation in wireless sensor networks can significantly reduce communication overhead compared to all the nodes sending their data to the base station. However, data aggregation complicates even more network security. This is due to the fact that every intermediate node could potentially modify, forge, or discard messages. Therefore, a single compromised node could be able to alter the final aggregation value. Intruder node and compromised node attacks are two major threats to security in sensor networks that use data aggregation techniques. An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

III. Key Expansion Algorithm

The key expand algorithm is used to expand the user supplied key to fill the expanded array S, so that S resembles an array of t=(2*r+4) random binary words determine by user supplied key K. it is differ from RC5 version where more words are derived from user supplied key K. These drive words are star in array S which are uses later encryption or decryption. in our proposed method we simplify the key expansion terms of RC6.

In first step:

key expansion is to copy the secret key K [0…….b-1] into array L[0……..c-1] this operation is done in natural manner, using u consecutive key bytes of K to fill up each successive word in L, in little endian order. The two magic constraints Pw and Q_w define for arbitrary were follows:

$$P_w = Odd((e- 2)2^w$$

$$Q_w = odd((Ø-1)2^w$$

These magic constant Pw and Q_w are uses for arithmetic progression modulo 2^w which provide randomness in table S.

In Second step:

In this step of key expansion we initialize array S to a particular pseudorandom bit pattern using an arithmetical progression modulo 2^w with magic constraint.

create and expanded key table S [0……2r+3] now we initialize this table by using magic constraints

$$S [o] = P_w$$

$$For\ i = 1\ to\ (2r+3)\ do$$

$$S [i] = S [I -1] + Q_w$$

In third step:

The third algorithm steps of key expansion are to mix in the user's secret key in three phrases over the array S and L. More precisely, due to the potentially different sizes of S and L, the larger array will be processed three times, and the other array may be handled more times.

Mix the secret key into table, S

```
I = j = 0 ; A = B = 0;
V = 3 x max { c, 2r + 4 }
For s = 1 to v do
{
 A = S [ i ] = S [ i ] + A + B) < < < 3
B = L [ j ] = ( L [ j ] + A + B ) < < < ( 3 + i )
J = ( j + 1 ) mod c
}
```

Key expansion function is an one way function so no one can determine secret key

Encryption :

This is a second phase of proposed scheme it composed with three states: pre-whitening, an inner

loop of rounds, and post-whitening. Pre-whitening and post-whitening remove the possibility of the plaintext revealing part of the input to the first round of encryption and the cipher text revealing part of the input to the last round of encryption.

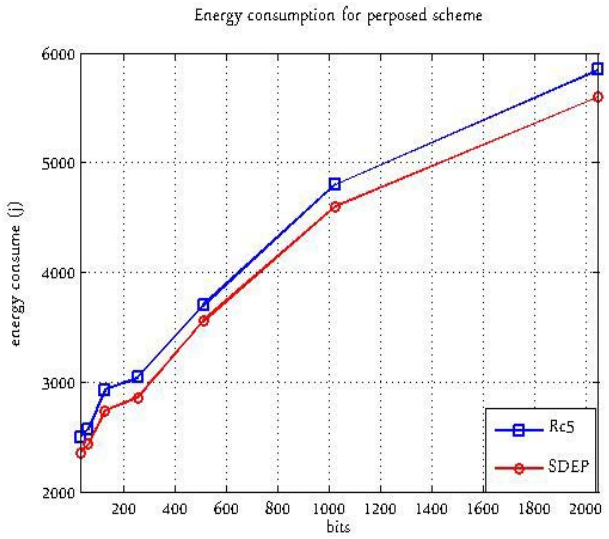
We uses four W bit input register A, B, C, D registers B and D undergo pre-whitening the register B and D put through the quadratic equation and rotated ($\log_2 w$) bits to the left respectively and these value store in variable t and u now register A XOR with t and left shift by u bits and added to round key S [2i].

Similarly C is XOR with the value of u and left shift by t bits. Now it added to round key S [2 i + 1]
for I = 1 to r do

{
T = (B * (2 B + 1)) < < < $\log_2$ W
U = (D * 2 (2 D + 1)) < < < $\log_2$ W
A = ((A XOR t) < < < u + S [2i]
C = ((C XOR u) < < < t) + S [2i + 1]
(A, B, C, D) = (B, C, D, A)
}

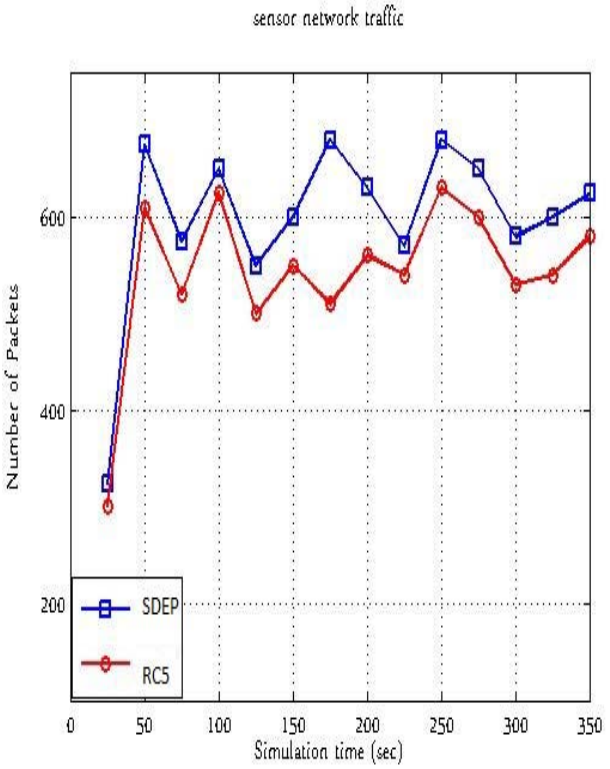
IV. Results

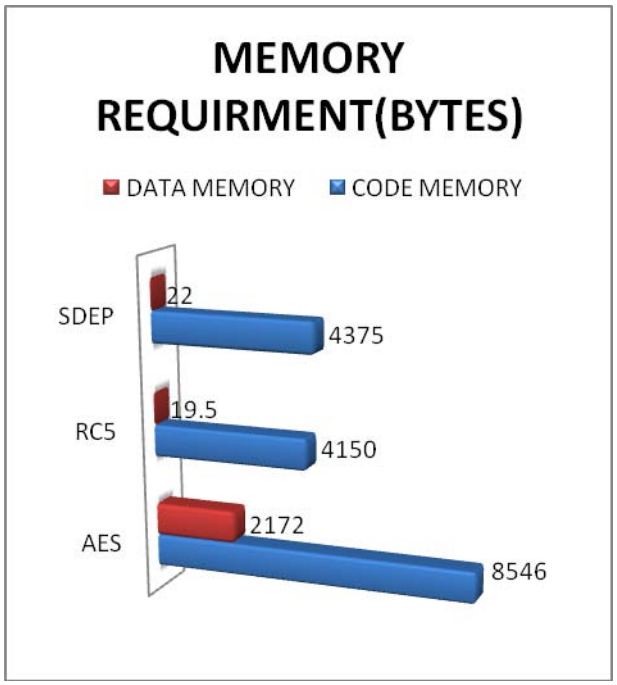
In this paper we have proposed a new algorithm for the security purpose in wireless sensor network. We also perform evaluation this new approach by comparing with two alternative popular algorithm AES and RC5. We investigate performance of this new algorithm based on memory requirements and the bandwidth according to our result the bandwidth for SDEP is much less than AES and RC5. Memory requirement for both code and data is less than AES and nearly equal to RC5. According to these simulation results our new algorithm SDEP much better than RC5 and AES in term of memory requirement, bandwidth requirement and time delay. It is also very much energy efficient.

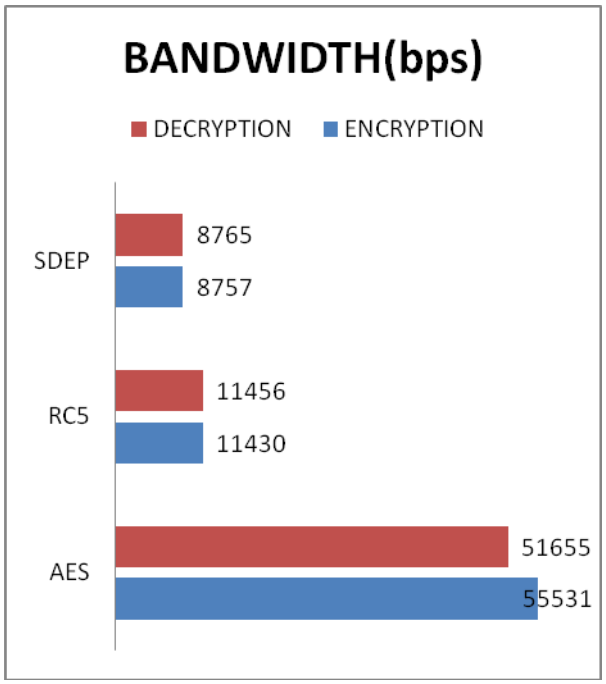


Sensor Network traffic

An Experiment was performed with 15 sources and a simulation time 350s. The maximum packets transmission of the sensor nodes could be achieved in 50 s. the number of packet transfer at that time are 680 in SDEP and 600 in RC5 so the delay time for SDEP is less than RC5 Figure2 show the estimated maximum transmission network packet for the sensor node.







V. Conclusion

This paper proposed a new security scheme for wireless sensor network in which we use the concept of RC6 for encryption and decryption of sensor data. In first modification of RC6 in key expansion step is static number of rotation and in encryption method we perform some function parallel based on RC 5 concept so it increase the throughput of SDEP. In this paper we also compare our new algorithm with AES and RC5 which show that proposed scheme is best useful in end to end encryption in wireless sensor network.

References Références Referencias

1. R. L. Rivest, "The RC5 Encryption Algorithm", in Proc. 1994 LeuvenWorkshop on Fast Software Encryption, 1995, pp. 86.96.
2. C. Karlof, N. Sastry, and D. Wagner. Tinysec: a link layer security architecture for wireless sensor networks. In SenSys 04: Proceedings of the 2nd international conference on Embedded networked sensor systems, pages 162.175, New York, NY, USA, 2004. ACM Press.
3. B. Kadri, A. Mhamed, and M. Feham "Secured Clustering Algorithm For Mobile Ad Hoc Networks", *International Journal of Computer Science and Network Security*, Vol.7, No.3, pp 27-34.2007.
4. J. Albath, and S. Madria. "Practical Algorithm for Data Security (PADS) in Wireless Sensor Networks". *MobiDE'07, , Beijing, China.2007*
5. Koo Woo K, Lee H, Kim Yong H, Lee Dong H (2008). Implementation and Analysis of New Lightweight Cryptographic Algorithm Suitable for Wireless Sensor Networks, Information Security and Assurance, ISA, pp. 73-76.
6. A. Perrig, R. Szewczyk, V. Wen, D. Culler, J. D. Tygar. SPINS: Security Protocols for Sensor Networks, in Wireless Networks Journal (WINE), September 2002.
7. P. Ganesan, R. Venugopalan, P. Peddabachagari, A. Dean, F Mueller, and M Sichitiu, "Analyzing and Modeling Encryption Overhead for Sensor Network Nodes", WSNA"03, September 19, 2003, San Diego, California,USA.
8. *Pawan Kumar Goel , Vinit Kumar Sharma ,* International Journal of Science Technology & Management , IJSTM Vol. 2, Issue 2, April 2011
9. Sencun Zhu, Sanjeev Setia, Sushil Jajodia. "LEAP: Efficient Security Mechanisms for Large Scale Distributed Sensor Networks", *In The Proceedings of the 10th ACM Conference on Computer and Communications Security*, 2003.
10. S. M. Hwang, and E. N. Huh. "An Efficient Topology Control and Dynamic Interval Scheduling Scheme for 6LoWPAN", in Gervasi et al. (Eds.): ICCSA, Part I, LNCS 5592, pp. 841–852. 2009
11. Anushiya A Kannan, Guoqiang Mao and Branka Vucetic, "Simulated Annealing based Wireless Sensor Network Localization", *Journal of Computers,* Vol. 1, No. 2, pp 15-22, May 2006.

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Automated Listing for Exportable Products

By Md.Sarwar Kamal, Sonia Farhana Nimmy & Mohd. Kamal Uddin
BGC Trust University Bangladesh, Chittagong

Abstract - Business plays a pivotal role for the development of human civilization as well as change of the status of the country in this whole universe. To compare and compute the export trend for any organization or a country is an indispensable for numerous intelligent systems to measure the national and international gain or loss for developing countries as well as for developed countries. The main purpose of this research is to develop a dynamic business prediction model for person, organization, Institute, Ministry of Commerce, Ministry of Finance ,Ministry of Economics, Prime Minister office and last but not least for whole world to predict the exact demand for exportable products and formulate export policy. Bangladesh Export Promotion Bureau (BEPB) helped us by providing the valuable data set and information. In this research activities we have had first classify the data by using Support Vector Machine (SVM), a latest data classification technique in the field of data processing. SVM reduce the redundant data from vast amount of data. After getting processed data we have used first-order logic (FL) to build a knowledgebase which will work as background knowledge for our computation. Finally Bayes' Network (BN) has used to perform the proper prediction by using the knowledge base information. Based on the result of BN, we have made a list of emerging products those have major impact on the prosperity of the country or organization.

Keywords : *SVM, Bayes' Network, Intelligent System, First order logic, Knowledge base, Export Trend.*

GJCST Classification: *H.2.8*

AUTOMATED LISTIN G FOR EXPORTABLE PRODUCTS

Strictly as per the compliance and regulations of:

Automated Listing for Exportable Products

Md.Sarwar Kamal [α], Sonia Farhana Nimmy [σ] & Mohd. Kamal Uddin [ρ]

Abstract - Business plays a pivotal role for the development of human civilization as well as change of the status of the country in this whole universe. To compare and compute the export trend for any organization or a country is an indispensable for numerous intelligent systems to measure the national and international gain or loss for developing countries as well as for developed countries. The main purpose of this research is to develop a dynamic business prediction model for person, organization, Institute, Ministry of Commerce, Ministry of Finance ,Ministry of Economics, Prime Minister office and last but not least for whole world to predict the exact demand for exportable products and formulate export policy. Bangladesh Export Promotion Bureau (BEPB) helped us by providing the valuable data set and information. In this research activities we have had first classify the data by using Support Vector Machine (SVM), a latest data classification technique in the field of data processing. SVM reduce the redundant data from vast amount of data. After getting processed data we have used first-order logic (FL) to build a knowledgebase which will work as background knowledge for our computation. Finally Bayes' Network (BN) has used to perform the proper prediction by using the knowledge base information. Based on the result of BN, we have made a list of emerging products those have major impact on the prosperity of the country or organization.

Keywords : SVM, Bayes' Network, Intelligent System, First order logic, Knowledge base, Export Trend.

I. Introduction

This paper aims to evaluate the export performance of Bangladesh in the context of rapid trade liberalization with its neighboring countries and thereby fit a future trend in its export business with an intention to emphasize on profitable goods in a systematic ranking of them. In the broadest sense, exporting refers to the sale of goods or services produced by a company based in one country to customers that reside in a different country. Export offers strategic benefits. Practically, the export businesses of Bangladesh are more or less dependent on a few products and selected markets. Most of the products have low income elasticity of demand. At the same time, it is also true that the country's export sector has undergone substantial structural changes during the last decades. A close study of this sector indicates that a significant shift occurred in this area from the erstwhile jute-centric export to ready-made garments centric one, with consequent dominance of the non-traditional manufactured commodities as against traditional primary ones.

Bangladesh, a developing and low per capita income generating country of South Asia, very often faces trade deficits in the context of international business (Rahman 2005). According to recent statistics of Bangladesh Bank, it is found that the trade deficit in goods goes beyond 500 crore dollars in the first seven month of current financial year 2011-2012. On the other hand, the trade deficit in service sector stands 160 crore 70 lac dollars for the same period. The service sector basically includes insurances, travels etc. Export growth in the first eight months of the current financial year declined to 13 per cent compared with 40.28 per cent during the same period in the past financial year as the ongoing economic crisis in Europe and United States continued to affect exports of jute goods and ready-made garments. Since this trade imbalance has some economic and political implications on the Bangladesh economy, Bangladesh considers it as a matter of great concern. Limited export base, backward industries, inadequate infrastructure, lower productivity, devaluation of money, higher tariffs, huge illegal trade are the main reasons of this trade imbalance.

Bangladesh has become less competitive. In the latest edition (September 2010) of Global Competitiveness report (GCR), Bangladesh ranks 108, a step down from what was shown in the 2010 report. The survey conducted by the World Economic Forum (WEF) among 142 countries, downgraded Bangladesh on three key problematic factors – inadequate power and energy infrastructures, frequent corruption and inefficient government bureaucracy.

According to Export Promotion Bureau, Bangladesh exports 173 products but its export is still mainly dependent on six products which contribute almost 88 per cent of the total export of which ready-made garments alone contributes about 76 per cent. The contribution of this products are woven garments 37.11 per cent, knit wear 40.01 per cent, froze food 2.73 per cent, jute goods 4.86 per cent, leather 1.40 percent and fertilizer and chemical products 1.26 per cent. The EPB source also informs that Bangladeshi products have markets in 186 countries with only four markets such as USA (26 per cent), EU (53 per cent), Canada (4 per cent), and Japan (2 per cent) having a contribution of about 85 per cent of country's export earnings. The other 182 markets contribute only 15 per cent. Among

Author α σ : Lecturer, Computer Science and Engineering, BGC Trust University Bangladesh, Chittagong.
Author ρ : Assistant Professor, Faculty of Business Administration, BGC Trust University Bangladesh, Chittagong.
E-mail α : sarwar.saubdcoxbazar@gmail.com
E-mail σ : nimmy_cu@yahoo.com
E-mail ρ : ksakib17@yahoo.com

the four major markets, two markets – USA and EU together contribute almost 79 per cent. Heavy dependence on these two markets is not a very comfortable situation for a developing country like Bangladesh and hence should seriously think about fixing a new export strategy for this sector.

This research activity organized as follows. In section 2 we have outlined the scope and importance of the study. In section 3 we have proposed our model of this work span. In section 4 we depict the Support Vector Machines (SVM) to classify the products. In section 5 we transmit Fuzzy C-Means Algorithm to clarify the small change of the export rate. According to the logic of SVM and FCM, in section 6 where we delineate our programmed pseudo code which worked to sort the products based on their amounts of export units. In next section we illuminate Knowledgebase and Bayes' Network (BN) to check the probability of the best exportable products. Last section we have used Fuzzy matching degrees to prepare the list of best fit of the exportable products. In section 9 we have had limned the result of the experiments.

II. Scope and Importance of the Study

This study is based on information and data collected through study and small sample survey. The sample survey was conducted on a small number of export firms dealing in ready-made garments, shrimps, leather, tea, shoes and ceramic items. Although a more comprehensive study was necessary to reveal the important dimensions of export business, the study had to be completed on a small scale due to paucity of finance and time. Yet it is expected that the findings will be representative because trends have a unique and ample shape from the macro point of view.

III. The Model

1. At first we have imposed Support Vector Machines (SVM) to classify the product from ambiguity.
2. Then we have concentrated on Fuzzy C-Means Classification so that the classification becomes more accurate.
3. Programmed according to the appropriate algorithm
4. Bayes Network and Knowledgebase to check the best products.
5. Fuzzy degree matching.
6. Result analysis.

IV. Support Vector Machine

In this work SVM made the whole work clean by enabling the proper classification of exportable products from any kinds of products list. Support Vector Machine (SVM) is one of the latest clustering techniques which enables machine learning concepts to amplify predictive accuracy in the case of axiomatically diverting data those are not fit properly. It uses inference space of linear functions in a high amplitude feature space, trained with a learning algorithm. It works by finding a hyperplane that linearly separates the training points, in a way such that each resulting subspace contains only points which are very similar. First and foremost idea behind Support Vector Machines (SVMs) is that it constituted by set of similar supervised learning. An unknown tuple is labeled with the group of the points that fall in the same subspace as the tuple. Earlier SVM was used for Natural Image processing System (NIPS) but now it becomes very popular is an active part of the machine learning research around the world. It is also being used for pattern classification and regression based applications. The foundations of Support Vector Machines (SVM) have been developed by V.Vapnik.

Two key elements in the implementation of SVM are the techniques of mathematical programming and kernel functions. The parameters are found by solving a quadratic programming problem with linear equality and inequality constraints; rather than by solving a non-convex, unconstrained optimization problem. The flexibility of kernel functions allows the SVM to search a wide variety of hypothesis spaces. All hypothesis space help to identify the Maximum Margin Hyperplane(MMH) which enables to classify the best and almost correct data The following figure shows the process of SVMs selection from large amount of SVMs.

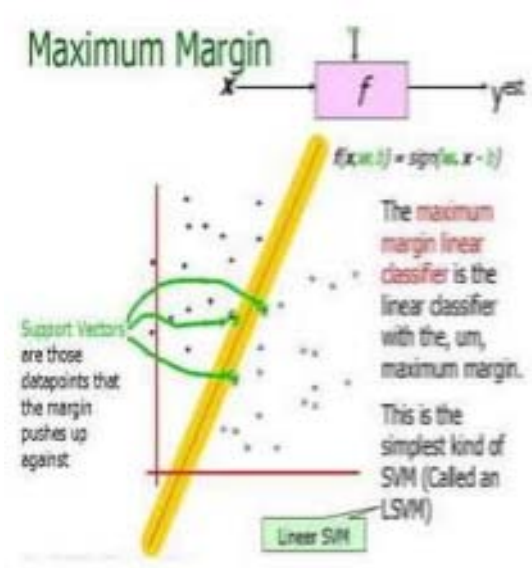


Fig 1: Selection of Support vectors s

Expression for Maximum margin is given as [4][8] (for more information visit [4]

$$\text{margin} \equiv \arg\min_{\mathbf{x}\in D} d(\mathbf{x}) = \arg\min_{\mathbf{x}\in D} \frac{|\mathbf{x}\cdot\mathbf{w}+b|}{\sqrt{\sum_{i=1}^{d} w_i^2}}$$

a) Euclidian Distance Measurement

To determines the exact maximum Margin Hyperplane it is very essential to measure the vertical

distances between any points and a straight line. These distances are the nothing but the Geometric distances from a line to a point. Suppose a straight line is 5x+3y+6=0 and we want to determine the distance of a coordinate A(3,2). According to the formula the Geometric distance of point a to the given line will be as follows:

Distance d = $\frac{5*3+3*2+6}{5*5+3*3}$ Generally we can say that the distance from ax+by+c=0 and a point $A(x_1,y_1)$ will be as follows: W= $\frac{ax1+by1+c}{a*a+b*b}$

While we will select the region of MMH the distance measurement equation plays an important role in this purpose. The following figure shows the Euclidian distances of the points A and the straight line ax +by +c=0

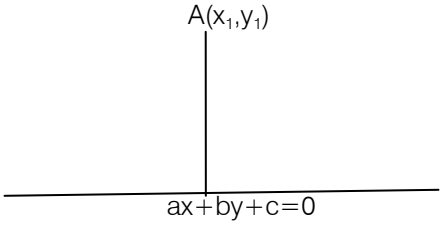


Fig 2 : Distance from a point A(x1,y1) and a straight line

b) Maximum Margin Hyperplane

The above illustration is the maximum linear classifier with the maximum range. In this context it is an example of a simple linear SVM classifier. Another interesting question is why maximum margin? There are some good explanations which include better empirical performance. Another reason is that even if we've made a small error in the location of the boundary this gives us least chance of causing a misclassification. The other advantage would be avoiding local minima and better classification. The goals of SVM are separating the data with hyper plane and extend this to non-linear boundaries using kernel trick [8] [11]. For calculating the SVM we see that the goal is to correctly classify all the data. For mathematical calculations we have,

[a] If Y_i= +1; $wx_i + b \geq 1$

[b] If Y_i= -1; $wx_i + b \leq 1$

[c] For all i; $y_i (w_i + b) \geq 1$

In this equation x is a vector point and w is weight and is also a vector. So to separate the data [a] should always be greater than zero. Among all possible hyper planes, SVM selects the one where the distance of hyper plane is as large as possible. If the training data is good and every test vector is located in radius r from training vector. Now if the chosen hyper plane is located at the farthest possible from the data [12]. This desired hyper plane which maximizes the margin also bisects the lines between closest points on convex hull of the two datasets. Thus we have [a], [b] & [c].

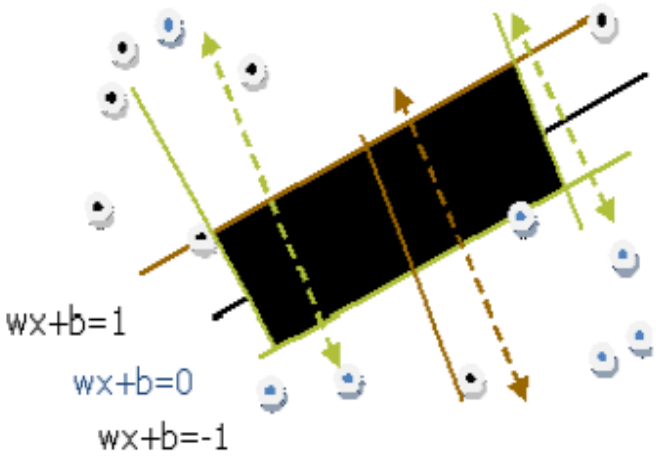


Figure 3 : Representation of Hyper planes. [9]

The figure above where the black shade area is the Maximum Margin Hyper Plane (MMH) which is the key data set area for experiment. Distance of closest point on hyperplane to origin can be found by maximizing the x as x is on the hyper plane. Similarly for the other side points we have a similar scenario. Thus solving and subtracting the two distances we get the summed distance from the separating hyperplane to nearest points. Maximum Margin = M = 2 / ||w||

c) Our Contribution

In this research we explore the concepts and technique of SVM to classify the data collected in our experiments. We have collected) data from Export Promotion Bureau of Bangladesh (EPBB) . To assess these large amounts of data we have found that SVM is very efficient and exact technique in our proceedings. By imposing the SVM, we have mapped the data to meaning full format to organize knowledgebase have shown in table 1 to 2.

V. Fuzzy c-Means Algorithm

Fuzzy C-means Algorithm capitalizes the data set to exact classes of given input under the value of membership function. Fuzzy C-means Algorithm is formulated as the minimization of the following objective function:

$$J_m(U,\mathbf{V}) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} D_{ik}^{2} \qquad (1)$$

Where, $U \in M_{fcn}$, $\mathbf{V} = (\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_c)$, $\mathbf{v}_i \in R^p$ is the i^{th} prototype m>1 is the fuzzifier and

$$D_{ik}^{2} = \|\mathbf{x}_i - \mathbf{v}_k\|^2$$

The objective is to find that U and **V** which minimize J_m

The Steps fro FCM Algorithm:

1. Choose: $1 < c < n$, $1 < m < \infty$, $\in$ = tolerance, max iteration = N
2. Calculation of membership values as according to the equation (2)

$$u_{ij} = \left[\sum_{k=1}^{c} \left(\frac{D_{ij}}{D_{ik}} \right)^{\frac{2}{m-1}} \right]^{-1} \forall i, j \qquad (2)$$

3. Computer the centroids values according to the equation (3)

$$\mathbf{v}_i = \left(\frac{\sum_{k=1}^{n} u_{ik}^{m} \mathbf{x}_k}{\sum_{k=1}^{n} u_{ik}^{m}} \right) \forall i \qquad (3)$$

4. Selection of new multiplier fields.
5. Repeat the step 2 until the algorithm has converged.

VI. Proposed Algorithmic Pseudo Code

In this part of the work we have made crucial steps to make sort the products based on the outcome units of export at various fiscal year. By using the knowledge of section 4 and 5 we have had mapped the following Pseudo code:

```
Define change(a, b) { char *ch=x[a];
X[a]=x[b]; x[b]=ch; }
Define d(i) x[i][depth]
void valuechange(int i, int j, int n, char *x[])
{while (n-- > 0) {change (i, j);i++;j++;}}
void short(char *x[], int n)
{ sort1(x, n, 0); }
void sort1(char *x[], int n, int depth)
{ int a, b, c, d, r, v; if (n <= 1) return; a = rand() % n;
change(0, a);v = d(0);a = b = 1;c = d = n-1;
for (;;) {
while (b <= c && (r = d(b)-v) <= 0) {if (r == 0) {
change(a, b); a++; }b++;}
while (b <= c && (r = d(c)-v) >= 0) {if (r == 0) {
change(c, d); d--; }c--;}if (b > c) break; change(b,
c);b++;c--;}r = min(a, b-a); valuechange(0, b-r, r, x); r
= min(d-c, n-d-1); valuechange(b, n-r, r, x);r = b-a;
sort1(x, r, depth);if (d(r) != 0)sort1(x + r, a + n-d-1,
depth+1);
r = d-c; sort1(x + n-r, r, depth);}
```

VII. Knowledge Base for Collected Data

A knowledge base in artificial intelligence is a place where information are stored or designed for machine or device by which it will work. In general, a knowledge base is a consolidate stock for information: a library, a database of related information about a particular subject could all be considered to be examples of knowledge bases. The process of building knowledge base is called knowledge engineering. A knowledge base is integrated collection of choosing logic, building a knowledge base, implementing [31] the proof theory, inferring new facts. The main advantage of engineering is that it requires less commitment and thus less work. To help the focus the development of knowledge base and to integrate the designer's thinking the following five step methodology can be used:

1. Decide what to talk about
2. Decide on a vocabulary of predicates, function, and constant.
3. Encode general knowledge about the domain.
4. Encode a description of the specific problem instance.
5. Pose queries to the inference procedure and answers.

In our work we have described a simple method of probabilistic inference that is, the computation from observed evidence of posterior probabilities for query propositions. We have used the joint probability as the knowledge base from which answer to all question may be derived. We have had built the knowledge base by considering two Boolean variables. The table 3 is an example of two valued propositional logic which is the bases of knowledge base representation:

	B		⌐ B	
	C	⌐ C	C	⌐ C
A	111	110	101	100
⌐ A	011	010	001	000

Table 1: Concepts of propositional logic to design a Knowledge Base using the proposition of Boolean events A, B and C

Based on table 1, we have designed the knowledge base (Joint probability distribution) for our research activity. Here we have considered those events which have true (one or 1) Boolean values. Table 2 is an example of knowledge base for events A, B and C:

	B		⌐B	
	C	⌐ C	C	⌐ C
A	P(A)*P(B) *P(C)	P(A)*P(B)* P(⌐ C)	P(A)*P(⌐ B)*P(C)	P(A)*P(⌐ B)*P(C)
¬A	P(⌐A)*P(B) *P(C)	P(⌐ A)*P(B) *P(⌐ C)	P(A)*P(⌐ B)*P(C)	P(A)*P(⌐ B)*P(C)

Table 2 : Fully Joint probability distribution

By keeping the similarities s with the table 2, we compared our factors as the products are exportable or not. The designing of knowledge base for the factors which we are considered are given in table 3:

	Exportable		⌐ Exportable	
	2009	2010	2009	2010
Tea	73/92*51/62*43/72 =0.125	83/102*21/52*23/52=0.044	30/53*21/32*10/39 =0.103	20/51*25/48*20/49 =0.037
⌐ tea	27/46*35/44*30/43 =0.279	10/29*12/43*35/48 =0.099	20/31*20/43*30/43 =0.231	10/41*20/43*30/43 =0.082

Where

0.125+0.044+0.103+0.037+0.279+0.099+0.231+0.082=1

VIII. Bayes'Theorem

Bayes' theorem and conditional probability are opposite to each other. Given two dependent events A and B. The conditional probability of P (A and B) or P (B/A) will be P (A and B)/P (A). Related to this formula a rule is developed by the English Presbyterian minister Thomas Bayes (1702-61).According to the Bayes rule it is possible to determine the various probabilities of the first event given the outcome of the second event in a sequence of two events.

The conditional probability:

$$P(B/A) = \frac{P(AandB)}{P(A)} \quad (1)$$

The equation (1) will help to find out the probabilities of B after being occurrences of the A. we get the Bayes' theorem for these two events as follows:

$$P(A/B) = \frac{P(A).P(B/A)}{P(B)} \quad (2)$$

If there are more events like A1, A2, and B1, B2.In this case the Bayes theorem to determine the probability of A1 based on B1will be as follows:

$$P(A1/B1) = \frac{P(A1).P(B1/A1)}{P(A1).P(B1/A1) + P(A2).P(B2/A2)}$$

Now applying the Bayes theorem on table 5 we have got the following outcomes:

If export condition is "Exportable" and year=2009 then P (Tea | Export condition="Exportable" ^Year=2009)=

P (Export condition="Exportable" ^Year=2009)= 0.125

P (Tea | Export condition="Exportable" ^Year=2009)=0.125+0.279 = 0.404

The total resultant of Bayes Theorem of all data considering Exportable or not are depicted following table 4:

Rule	Outcome
P(Tea|Export condition="Exportable" ^year=2009)	90.9%
P(Tea|Export condition="Exportable" ^year=2010)	90.8%
P(Spices|Export condition="Exportable" ^year=2009)	80.8%
P(Spices|Export condition="Exportable" ^year=2010)	81.1%
P(Veg|Export condition="Exportable" ^year=2009)	79.1%
P(Veg|Export condition="Exportable" ^year=2010)	79.9%
P(tobacco|Export condition="Exportable" ^year=2009)	69.2%
P(tobacco|Export condition="Exportable" ^year=2009)	68.9%

IX. Fuzzy Matching Degrees for Comparing Marginal Change

The operational block of the instance matching integrates ontology alignment, retrieves semantic link clouds of an instance in ontology and measures the terminological and structural similarities to produce matched instance pairs. Pseudo code of the Instance Matching algorithm

Algo. InstanceMatch (ABox ab1, ABox ab2, Alignment A)
for each insi element of ab1
cloudi=makeCloud(insi,ab1)
for each insj element of ab2
cloudj=makeCloud(insj,ab2)
if $\forall$ a(c1; c2) elements of A|c1 elements of Block(ins1:type) ^
c2 elements of Block(ins2:type)
if Simstruct(cloudi; cloudj) $\geq \delta$
imatch=imatch $\bigcup$ makeAlign(insi; insj)

Let us consider the fuzzy matching for the mixing products s on the input data set .The degree to which the input target data set satisfy the conditions of fuzzy rules and conditions .Suppose Product 1 is defined by rules R_1 and Product 2 is defined by rules R_2.In this case the matching degree will be represented by as follows:

Matching Degree (IMAGE X,R_1)=μ(IMAGE X)
Matching Degree (IMAGE Y,R_2) =μ (IMAGE Y)
Where μ is the fuzzy membership function.

The fuzzy matching determines the actual outcome for fuzzy optimization which is accomplished here by fuzzy matrix. Here is a graphical view of fuzzy matching degree for IMAGE Y as follows:

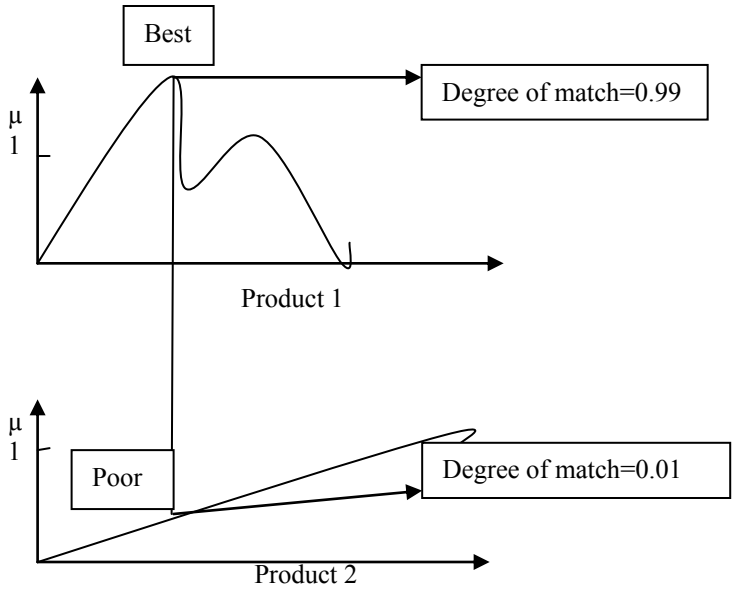


Fig 2 : The fuzzy matching degree for various products

X. Experimental Result

Based on the fuzzy matching degree variation we have noted the following two tables for Agricultural and Engineering products which are portrayed at table 5 and 6.The whole result of any given data set will be same as table 5 and 6 for this research. That why here we briefly characterized these two table as the sample of any amount of data set.

Agricultural products

Products Names	Amounts	Fuzzy Degrees	Remarks
Tea	<<1000 Units	0.98	Very Good
Spices	<1000 units	0.9	Good
Vegetables	1000 units	0.8	Good
tobacco	>1000 units	0.75	Satisfactory good
Cut flower	>900 units	0.70	Average
Fruits	>800 units	0.68	Poor
Dry food	>700 units	0.64	Very poor

Table 5 : The Resultant table for Agriculture products

Engineering products

Products Names	Amounts	Fuzzy Degrees	Remarks
Engineering Equipment	375 Units	0.92	Very Good
Electric Products	312 units	0.84	Good
Bicycle	210 units	0.75	Good
Iron Steal	157 units	0.71	Satisfactory good
Copper Wire	142 units	0.67	Average
Stainless Steel ware	120 units	0.57	Poor
Iron Steel	792 units	0.635	Very poor

Table 6 : The Resultant table for Engineering products

XI. Conclusion

The trade deficit of Bangladesh in international business is historic and recently it is growing very sharply. The government of Bangladesh must try to remove all structural impediments of export expansion. Maximum efforts are to be devoted to diversify the export base. To encourage capacity building and to

improve productivity and competitiveness in terms of both quality and cost, domestic and foreign investments must be attracted through appropriate government policies. Availability of sound physical and economic infrastructure, superior product quality, sufficient institutional facilities for banking, credit and insurance, improved law and order situation, labor unrest free environment, a honest and efficient administration and continuous political stability must be ensured for the country's export promotion.

References Références Referencias

1. Bhattacharjee M. Kumar & Taher M. A, "Environmental opportunities and Challenges for Industrial Investment: Bangladesh as Case". *The Chittagong University Journal of Business Administration,* Vol. 18, p. 171-182, 2003
2. Chowdhury A. J. M. Nuruddin & Islam S. M. Nazrul, "Handling Risks in Export Marketing Operations - A Study of Some Export Firms Operating in Chittagong". *The Chittagong University Journal of Commerce*, Vol. 15, p. 1-18, 2000
3. Rahman M. Mafizur, "Bangladesh-India Bilateral Trade: Cases and Measures for Improvement", Discipline of Economics, University of Sydney, NSW 2006, Australia. June 2005
4. Burges C., "A tutorial on support vector machines for pattern recognition", In "Data Mining and Knowledge Discovery". Kluwer Academic Publishers, Boston, 1998, (Volume 2).
5. Zahid. H. "Constraints of Export Diversification", *The Daily Star*, Dhaka, March 24, 2011, p.46
6. Recent Report of Export Promotion Bureau, Bangladesh Government of the People's Republic of Bangladesh, Ministry of Industries, *Industrial Policy* July 1999.
7. Rashid. H, "Barriers hampers export to India*", The Daily Star,* Dhaka, April 27, 2011, p.15
8. Nello Cristianini and John Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods", Cambridge University Press, 2000
9. Image found on the web search for learning and generalization in svm following links given in the book above.
10. Guesgen, H.;Hertzberg, J.; Lobb, R.;Mantler, A. { Bu_ering Fuzzy Maps in GIS, Spatial Cognition and Computation, 2003
11. Tom Mitchell, Machine Learning, McGraw-Hill Computer science series, 1997.
12. J.P.Lewis, Tutorial on SVM, CGIT Lab, USC, 2004.

Global Journal of Computer Science and Technology Volume XII Issue IX Version I April 2012
34

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Optimization and Efficient Transmission Schedule for 802.11 Networks with Jamming Characteristics

By Mrs.N.Elamathi & Dr.C.Chandrasekar
Trinity College for Women Periyar University, Namakkal

Abstract - In Wireless 802.11 networks, Multiple-path source routing allows data source node to distribute the total traffic among the possible available paths. However, in this case jamming effects were not considered. Recent work has presented jamming mitigation scheme, anti-jamming Reinforcement System on 802.11 networks by assessing physical-layer functions such as rate adaptation and power control. Rate adaptation algorithms significantly degrade network performance. Appropriate tuning of carrier sensing threshold allows transmitter to send packets even on jam that enable receiver to capture desired signal. Efficient schedules need to be investigated for redundant transmission to perform well in presence of jammer.

In this paper, the proposal in our work presents an Efficient Time and Transmission Schedule Scheme for wireless 802.11 networks in presence of jamming that guarantee low waiting time and low staleness of data. Schedules are optimal even jamming signal has energy limitations. Each packet is encoded by an error-correcting code (Reed-Solomon). Reed solomon code allow schedule to minimize waiting time of the clients and staleness of the received data. Jammers have restrictions on length of jamming pulses and length of intervals between subsequent jamming pulses.

Keywords : *Transmission Schedule, 802.11 Network, Jamming characteristics.*

GJCST Classification: *C.2.1,C.2.2*

Strictly as per the compliance and regulations of:

Optimization and Efficient Transmission Schedule for 802.11 Networks with Jamming Characteristics

Mrs.N.Elamathi [α] & Dr.C.Chandrasekar [σ]

Abstract - In Wireless 802.11 networks, Multiple-path source routing allows data source node to distribute the total traffic among the possible available paths. However, in this case jamming effects were not considered. Recent work has presented jamming mitigation scheme, anti-jamming Reinforcement System on 802.11 networks by assessing physical-layer functions such as rate adaptation and power control. Rate adaptation algorithms significantly degrade network performance. Appropriate tuning of carrier sensing threshold allows transmitter to send packets even on jam that enable receiver to capture desired signal. Efficient schedules need to be investigated for redundant transmission to perform well in presence of jammer.

In this paper, the proposal in our work presents an Efficient Time and Transmission Schedule Scheme for wireless 802.11 networks in presence of jamming that guarantee low waiting time and low staleness of data. Schedules are optimal even jamming signal has energy limitations. Each packet is encoded by an error-correcting code (Reed-Solomon). Reed solomon code allow schedule to minimize waiting time of the clients and staleness of the received data. Jammers have restrictions on length of jamming pulses and length of intervals between subsequent jamming pulses.

Keywords : Transmission Schedule, 802.11 Network, Jamming characteristics.

I. Introduction

Wireless 802.11 networks were designed under the assumption that all nodes are interested in transmission of data, and follow the rules of the protocol regarding when to send and when to permit other nodes to send. Jamming point-to-point transmissions in wireless 802.11 network can produce negative effects on data transport through the network. The effects of jamming at the physical layer provides an effective denial-of-service. The simplest methods to avoid jamming at the physical link layer is to provide solutions such as spread-spectrum or beaforming. It forces the jammers to extend a greater resource to reach at the desired goal. Spread-spectrum techniques are methods by which a signal (e.g. an electrical, electromagnetic, or acoustic signal) generated in a particular bandwidth is deliberately spread in the frequency domain, resulting in a signal with a wider bandwidth. These techniques are used for a variety of reasons, including the establishment of secure communications, increasing resistance to natural interference noise and jamming, to prevent detection, and to limit power flux density.

However, recent work has proposed that intelligent jammers can create cross layer protocol information to jamming attacks which in turn reduces the resource expenditure. It targets certain link layers and MAC representations and also link layers error detection and error correction protocols. So more sophisticaed anti-jamming traffic methods have proposed in the higher layer to name a few are channel surfing. Channel surfing is the practice of quickly scanning through different television channels r radio frequencies in order to find something interesting to watch or listen to. Modern viewers, who may have cable or satellite services beaming down dozens if not hundreds of channels, are frequently caught channel surfing. It is common for people to scan channels when commercial broadcasters switch from a show over to running advertisements.

The majority of anti-jamming techniques considers mainly diversity. It may employ multiple frequency bands, different MAC channels or multiple routing paths. To make effective use of this routing diversity, each source node must be able to make an intelligent allocation of traffic across the different available paths while also considering the potential effect of jamming on the resulting data throughput. Anyhow, the jamming at each network node depends on a number of unknown parameters like the strategy used by the individual jammers and the relative location of the jammers with respect to each transmitter-receiver pair.

In this paper, we introduce an efficient Time and Transmission schedules for wireless 802.11 networks in presence of jamming that guarantee low waiting time and low staleness of data. Proposed Time and Transmission schedule scheme is optimal even jamming signal has energy limitations. Each packet is encoded by an error-correcting code (Reed-Solomon). Reed solomon code allow schedule to minimize waiting time of the clients and staleness of the received data. Jammers have restrictions on length of jamming pulses

Author α : Research scholar Lecturer, Department of Computer Science Trinity College for Women Namakkal-1.
E-mail : emathi@yahoo.com
Author σ : Associate professor Department of Computer Science Periyar University Salem-11. E-mail : ccsekar@gmail.com

and length of intervals between subsequent jamming pulses.

II. Literature Review

The Widespread proliferation of IEEE 802.11 wireless networks makes them an attractive target for saboteurs with jamming devices [1]. They provide network access for both mesh and conventional clients. The focus of designing reliable UWA networks that is capable of transferring data from a variety of sensors to on-shore facilities [2]. Major impediments to the design of such networks were considered, which are: 1) severe power limitations imposed by battery power; 2) severe bandwidth limitations; and 3) channel characteristics including long propagation times, multi path, and signal fading.

In [3], investigated the problem of denial of service against data packets (e.g., IP packets) transmitted over WLAN protocols (i.e., IEEE802.11 and Bluetooth). It is easy to jam such communications at an energy cost [4] that is much lower than the transmitter's cost. Such attacks cannot only prevent communication within large areas [5] for long periods of time but can also lead to other more elaborate and coordinated attacks such as partitioning of a multihop ad hoc network or forcing packets to be routed over chosen paths [6].

A different defense strategy [7] involves sensors trying to out-compete the jammer by employing error correcting codes and increasing the node transmission power. Both evasion and competition strategies [8] are at an early stage of investigation by the community [9], and as these techniques mature an important area for study was understood and classifying the scenarios where one defense strategy is advantageous over another.

In [10], jamming-aware source routing traffic allocation methods in wireless mesh network. It maps lossy network flow optimization algorithm [11] to the asset allocation algorithm with the help of portfolio selection theory. Using the distributed algorithm multi-source multi-path optimal traffic allocation was computed for the intended source nodes. Distributed algorithm was precisely based on the decomposition in network utility maximization. The existing work also allows individual network nodes to locally characterize the jamming impact for the aggregate of source nodes [12]. As there is an uncertainty in achievable traffic rates portfolio selection theory allows data sources to balance the expected data throughput with the available value.

III. Efficient Time and Transmission Schedule for 802.11 Networks With Jamming Characteristics

The proposal work presents Efficient Time and Transmission Schedule for 802.11 Networks with Jamming Characteristics. The characteristics of jamming traffic are measured. The impact of resource utilization on jamming traffic is also identified. For all the resources, transmission schedule is assigned. The data source node for optimal multi-path traffic allocation is identified and localized. The use of scheduled resource utilization improves the data delivery rate of the source nodes. Simulations are carried out for the Time and Transmission Schedule for 802.11 Networks which is made on multiple traffic rate variances. Jamming characteristics for the resource utilization at the local source nodes are also evaluated.

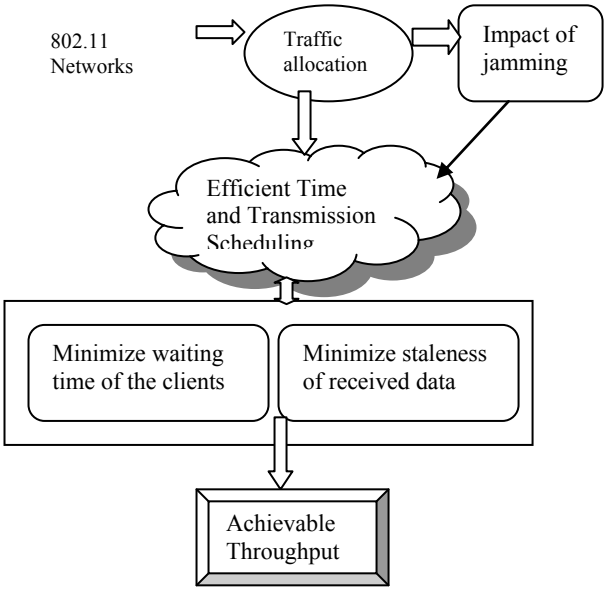


Figure 3.1: Architecture diagram of Efficient Time and Transmission Scheduling Scheme

The traffic is distributed among the available paths in the network. It also involves empirical jamming statistics. In result the impact of jamming is felt in the distribution of traffic network. 802.11 Networks are applicable to different scenarios. The job of source node is to make the allocation in an intelligent manner. It is also made available across different paths. The distribution of traffic provides with multiple source nodes for jamming and with multiple routing paths. The impact of jamming is characterized in the distribution of traffic. Proposed Time and Transmission schedule scheme (as shown in fig 1) is optimal even jamming signal has energy limitations. Each packet is encoded by an error-correcting code (Reed-Solomon). Reed solomon code allow schedule to minimize waiting time of the clients and staleness of the received data. Quadratic term is expressed in throughput due to the uncertainty involved in estimate. While formulating the multi-path traffic allocation risk-aversion parameter is the main factor considered.

a) Jamming Characteristic

The jamming model must be accurate enough to capture the characteristics of practical jammers, and,

at the same time, be simple enough for the optimization of network protocols. It has been recognized that the power supply is the most important limitation for the majority of practical jammers. A typical jammer is powered by a battery, which can be recharged from an external source, such as a solar cell array. The source node determines its own traffic allocation with the help of minimal message passing between sources. The goal of the jammer is to disrupt the normal operation of the broadcast system, which results in high waiting time and excessive power consumption of the clients. To that end, the jammer sends active signals over the channel that interfere with the signal sent by the server A set of sources with estimated parameters compensate on the presence of jamming on network traffic flow in distributed formulation for jamming-aware traffic allocation.

b) Efficient Time and Transmission Schedule

Multiple resources are allocated for difference MAC channels, multiple routing paths and multiple frequency bands. Risk-aversion is achieved for multiple resources. Resources are allocated to less risky paths than highly risky paths. Trade-off is maintained between expected throughput and estimation variance. It also varies with time or for various types of data. The data is delivered in the form of packets, each packet captures the current state of the information source. We assume that each packet includes exactly *P* information symbols. We also assume that transmission of P symbols of over the channel requires one unit of time.

We enumerate the packets, according to the time of their transmission. Each packet is encoded into a message that contains at least P symbols by using an error-correcting code, such as a Reed-Solomon [16]. The encoding ensures that any P symbols of the message are sufficient in order to reconstruct the original message.

A schedule is a sequence $\{t_1, t_2, \ldots . t_m\} \geq 1$, such that t_m is the amount of time required to transmit message m. Note that the length of message m is equal to $t_m P$.

A schedule $\{t_1, t_2 \ldots\}$ can also be defined by its transmission sequence $\{TS_1, TS_2 \ldots\}$, where TS_m represents the starting time of the transmission of message m, i.e., $TS_1 = 0$ and $TS_m = \sum_{n=1}^{m-1} t_n$ for $m > 1.$ (1)

In the first schedule, each encoded message contains $t.P$ symbols. Thus, the schedule transmits each message is transmitted over an interval of t time units and generates a new packet at times $0, t, 2t \ldots .$.

A wireless client begins to listen to the wireless channel upon a request for new information. In order to satisfy the request, the client must receive at least P symbols from the currently transmitted message. If the client fails to receive P symbols from the current message, it continue to listen to the channel, until it receives at least P symbols from one of the subsequent messages.

They are two key performance characteristics of the schedule: the expected waiting time and the maximum staleness of the received data.

Waiting time (S)): Let S be a broadcast schedule. Suppose that the client's request was placed at time t. Let k be the number of the message currently transmitted over the channel. Let t' be the first time the client receives at least P symbols from a message $k', k' \geq k$. Then, the waiting time of the client is defined as

$$WaitingTime(S) = t' - t \quad (2)$$

Following [15], [12], and [18], we assume that the clients' requests are distributed uniformly over time. Accordingly, the *Expected Waiting Time* of the clients is defined as follows:

$$\text{Expected Waiting Time (S)} = \lim_{w \to \infty} \frac{1}{w} \int_0^w WaitingTime(S) dw \quad (3)$$

The waiting time is an extremely important parameter for many time-sensitive applications. In addition, it is closely related to the amount of power spent by the client to obtain the information. The staleness of the data is defined to be the amount of time that passes from the moment the information is generated until it is delivered to the client. The staleness captures the quality of delivered information, because in dynamic settings the information becomes less and less relevant with time.

Staleness(S): Suppose that the client's request was placed at time t. Let k be the number of the message currently transmitted over the channel. Further, let $k' \geq k$ be the first message for which the client receives at least p symbols. Then, the staleness of the data is defined to be

$$Staleness(S) = t_{k'} - t \quad (4)$$

Suppose that a client arrives at time ts. The number of symbols received by the client from the currently transmitted message is equal to

$$k_t = ([\frac{ts}{t}]t - ts)P \quad (5)$$

If $k_t \geq P$, then the client will be able to decode this message, hence its waiting time is zero. Otherwise,

the client needs to wait for the next message, hence its waiting time is nt. It is easy to verify that if the clients are distributed uniformly over time, the expected waiting time is

$$\frac{P}{2k} = \frac{1}{2t} \quad (6)$$

While redundant transmission improves the expected waiting time of a schedule, it comes at a price in terms of the staleness of the received data. Indeed, if $k_t \geq P$, then the packet received by the client at time *ts*, was generated in time $\left\lfloor \frac{ts}{t} \right\rfloor t$, hence the staleness of the data is

$$ts - \left\lfloor \frac{ts}{t} \right\rfloor t \quad (7)$$

On the other hand, if $k_t < P$, then the client will get a new packet, hence the staleness is zero.

The example demonstrates that there exists a certain trade-off between waiting time and staleness in data broadcast systems. While finding a schedule that has minimum waiting time subject to a staleness constraint in a not-jammed channel is a relatively easy task, this task is much more complicated in the presence of a jammer.

c) Algorithm for Time and Transmission Schedule

Pseudo code of the Time and Transmission Scheduling algorithm is shown in below.

Input: Packets includes exactly *P* information symbols
Output: Transmission schedule S
t ← Time slot when the 1st packet arrives
while (unscheduled transmission) do
T← set of transmissions that are ready at slot t
Channel ← 0 **for** each T
i. T_w ← Calculate Waiting Time from eqn (2)
ii. T_{EW} ← Calculate Expected Waiting Time from eqn (6)
iii. D_s ← Calculate Staleness data from eqn (7)
while (($T_{EW,}$ D_s) = = low) do
Tr ← Transmission with low Waiting Time and Staleness data
Tr' ← Transmission with shortest deadline in Tr
if (Tr' misses deadline) **then** return unschedulable
S (ch) ← Tr'
ch ← ch + 1
end
t ← t + 1
end

As shown in the pseudo code, Input is taken as Packets includes exactly *P* information symbols and Outputs the Transmission schedule S. For any transmission, the time time slot t is assigned. The set of Unscheduled transmission T is taken with the time slot t. Initially, Channel offset ch is assigned to zero. For each Transmission T, Waiting Time, Expected Waiting Time and Staleness of data is calculated by using the equation (2), (6) and (7) respectively. If expected Waiting Time and Staleness of data is low then while loop is executed. For any transmission T, while scheduling at some slot t, if the algorithm determines that Tr' (deadline of Tr) is already less than t, then the algorithm terminates as it has failed to find a feasible solution. Otherwise, Tr' is assigned slot t and channel offset ch, the schedule is recorded as S[ch] = Tr' which is an final output.

IV. Performance Evaluation on Efficient Time and Transmission Schedule Scheme

In this section, we simulate various aspects of the proposed Time and Transmission scheduling algorithm, we evaluate the results of simulation using NS-2. For this purpose, we compared our algorithm with the ARES. The proposed NIRA using IRDM is evaluated in an efficient manner using NS2 simulator. Initially the experiment is evaluated with 100 nodes in a flat area of 100 * 100 m2. The nodes' incoming time (sec) is noted as t1, t2….tn. The routing discovery mechanism is taken for routing information to identify the route path from source to destination. The simulation results show that it takes 850 secs to transmit the packet from source to destination by choosing the path efficiently. The compared aspects were

i. Achievable Throughput
ii. Traffic allocation
iii. Execution time

Achievable Throughput: Average rate of successful message delivery over a communication channel. This data may be delivered over a physical or logical link, or pass through a certain network node

Traffic allocation: Transmission schedule is measured as percentage of cases the algorithm is able to find a feasible schedule among the total number of cases considered. The maximum number of traffic allocated at a field device at any point of time during the schedule.

Execution time: This is the total time required to create a schedule for all packets generated within T (least common multiple of periods) time slots.

V. Results and Discussion on Efficient Time and Transmission Schedule Scheme

Simulation works are carried out with NS2 with achievable throughput using traffic allocation problem

based on the resulting data throughput. The available paths are characterized based on throughput. The traffic rates incurred are non-negative and also define convex space. Due to the jamming characteristics traffic rate is reduced in the receiving node which again imposes stochastic constraint. The delay time is bound in traffic distribution algorithm. The traffic variance is computed during regular intervals. Exponential Weighted Moving Average (EWMA) method is applied to update traffic variance. Bandwidth utilization is low bound without loss of bandwidth. The performance metrics measured are Achievable Throughput, Traffic allocation and Execution time.

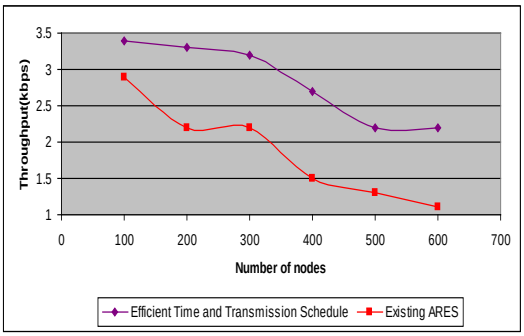


Figure 5.1: Throughput

Network throughput is the average rate of successful message delivery over a communication channel. This data may be delivered over a physical or logical link, or pass through a certain network node. The throughput is usually measured in bits per second (bit/s or bps), and sometimes in data packets per second or data packets per time slot. The figure 1 shows the output of the simulation by varying the number of nodes with in Wireless 802.11 networks. As the number of nodes increases, throughput increases. By comparing it with existing Jamming-aware source routing traffic allocation model, our Priority based resource utilization scheme is effective.

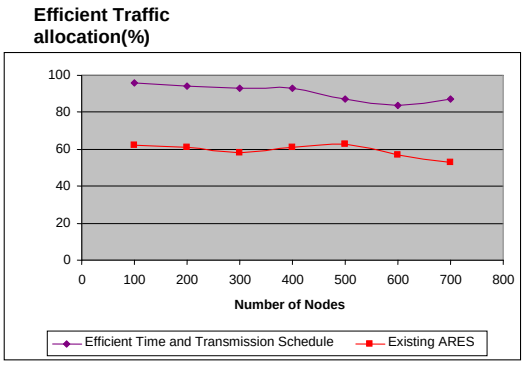


Figure 5.2 : Efficient traffic allocation

Figure 5.2 depicts the resultant graph of efficient traffic allocation on Proposed Time and Transmission Schedule and Existing ARES. For each simulation nodes can be changed from 100,200…. 700. Traffic allocation ratio is measured for each simulation. If number of nodes will be increased, Traffic allocation ratio gets automatically decreased in both Proposed Time and Transmission Schedule and Existing ARES. For example, In Proposed Time and Transmission Schedule scheme, nodes from 100 to 400, Traffic allocation ratio is 93 to 96 % and nodes from 500 to 700 traffic allocation ratio gets decreased to 84%. The above performance graph shows the Proposed Time and Transmission Schedule scheme outperforms well compared with ARES.

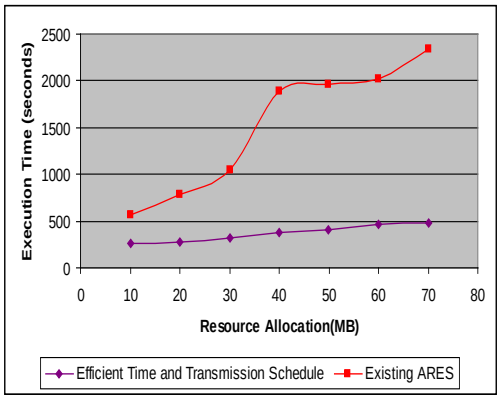


Figure 5.3 : Execution Time

The execution time of the Proposed Time and Transmission Schedule scheme and Existing ARES increases sharply with the increase of workload increases. Resource is allocated from 10 MB to 70 MB. Figure 5.3 shows that the execution time of Proposed Time and Transmission Schedule scheme is relatively low compared with ARES.

The simulation results show that our proposed Time and Transmission Scheduling scheme performs well in terms of Achievable Throughput, Traffic allocation and Execution time.

VI. Conclusion

In this paper, we have implemented an Efficient Time and Transmission schedule scheme for wireless 802.11 networks in presence of jamming. Proposed Time and Transmission schedule scheme guarantee low waiting time and low staleness of data. Schedules were optimal even jamming signal has energy limitations. Reed-Solomon error-correcting code has been used to encode an each packet while transmission that allow schedule to minimize waiting time of the clients and staleness of the received data. Jammers have restrictions on length of jamming pulses and length of intervals between subsequent jamming pulses. Experimental simulations were conducted to evaluate our Time and Transmission schedule algorithm. Simulation results show that our proposed Time and Transmission Scheduling scheme performs well.

References Références Referencias

1. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A survey," Computer Networks, vol. 47, no. 4, pp. 445–487, Mar. 2005.
2. E. M. Sozer, M. Stojanovic, and J. G. Proakis, "Underwater acoustic networks," IEEE Journal of Oceanic Engineering, vol. 25, no. 1, pp. 72–83, Jan. 2000.
3. J. Bellardo and S. Savage, "802.11 denial-of-service attacks: Real vulnerabilities and practical solutions," in Proc. USENIX Security Symposium, Washington, DC, Aug. 2003, pp. 15–28.
4. A. D. Wood and J. A. Stankovic, "Denial of service in sensor networks," IEEE Computer, vol. 35, no. 10, pp. 54–62, Oct. 2002.
5. G. Lin and G. Noubir, "On link layer denial of service in data wireless LANs," Wireless Communications and Mobile Computing, vol. 5, no. 3, pp. 273–284, May 2005.
6. W. Xu, K. Ma, W. Trappe, and Y. Zhang, "Jamming sensor networks: Attack and defense strategies," IEEE Network, vol. 20, no. 3, pp. 41–47, May/Jun. 2006.
7. D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," IEEE Journal on Selected Areas in Communications, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.
8. R. Leung, J. Liu, E. Poon, A.-L. C. Chan, and B. Li, "MP-DSR: A QoS aware multi-path dynamic source routing protocol for wireless ad-hoc networks," in Proc. 26th Annual IEEE Conference on Local Computer Networks (LCN'01), Tampa, FL, USA, Nov. 2001, pp. 132–141.
9. H. Markowitz, "Portfolio selection," The Journal of Finance, vol. 7, no. 1, pp. 77–92, Mar. 1952.
10. S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge,2004.
11. D. J. Thuente and M. Acharya, "Intelligent jamming in wireless networks with applications to 802.11b and other networks," in Proc. 25th IEEE Communications Society Military Communications Conference (MILCOM'06), Washington, DC, Oct. 2006, pp. 1–7.
12. D. B. Johnson, D. A. Maltz, and J. Broch, DSR: The Dynamic Source Routing Protocol for Multihop Wireless Ad Hoc Networks. Addison- Wesley, 2001, ch.5, pp. 139–172.
13. http://www.isi.edu/nsnam/n

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Advanced Methods to Improve Performance of K-Means Algorithm: A Review

By Ritu Yadav & Anuradha Sharma
Guru Jambheshwar University of Science & Technology,Hissar,Haryana India

Abstract - Clustering is an unsupervised classification that is the partitioning of a data set in a set of meaningful subsets. Each object in dataset shares some common property- often proximity according to some defined distance measure. Among various types of clustering techniques, K-Means is one of the most popular algorithms. The objective of K-means algorithm is to make the distances of objects in the same cluster as small as possible. Algorithms, systems and frameworks that address clustering challenges have been more elaborated over the past years. In this review paper, we present the K-Means algorithm and its improved techniques.

Keywords : *classification, clustering, k-means clustering, partitioning clustering.*

GJCST Classification: *F.2,F.2.m*

Strictly as per the compliance and regulations of:

Advanced Methods to Improve Performance of K-Means Algorithm: A Review

Ritu Yadav[α] & Anuradha Sharma[σ]

Abstract - Clustering is an unsupervised classification that is the partitioning of a data set in a set of meaningful subsets. Each object in dataset shares some common property- often proximity according to some defined distance measure. Among various types of clustering techniques, K-Means is one of the most popular algorithms. The objective of K-means algorithm is to make the distances of objects in the same cluster as small as possible. Algorithms, systems and frameworks that address clustering challenges have been more elaborated over the past years. In this review paper, we present the K-Means algorithm and its improved techniques.

Keywords : classification, clustering, k-means clustering, partitioning clustering.

I. Introduction

Clustering is a type of categorization imposed rules on a group of data points or objects. A broad definition of clustering could be "the process of categorizing a finite number of data points into groups where all members in the group are similar in some manner". As a result, a cluster is a aggregation of objects. All data points in the same cluster have common properties (e.g. distance) which are different to the data points laying in other clusters.

Cluster analysis is an iterated process of knowledge discovery and it is a a multivariate statistical technique which identifies groupings of the data objects based on the inter-object similarities computed by a chosen distance metric .Clustering algorithms can be classified into two categories: Hierarchical clustering and Partitional clustering [1]. The partitional clustering algorithms, which differ from the hierarchical clustering algorithms, are usually to create some sets of clusters at start and partition the data into similar groups after each iteration. Partitional clustering is more used than hierarchical clustering because the dataset can be divided into more than two subgroups in a single step but for hierarchy method, always merge or divide into 2 subgroups, and don't need to complete the dendrogram[2].

Cluster analysis of data is an important task in knowledge discovery and data mining. Cluster analysis aims to group data on the basis of similarities and dissimilarities among the data elements. The process can be performed in a supervised, semi-supervised or unsupervised manner. Different algorithms have been proposed which take into account the nature of the data and the input parameters in order to partition the data. Data vectors are clustered around centroid vectors. The cluster the data vector belongs to is determined by its distance to the centroid vector. Depending on the nature of the algorithm, the numbers of centroids are either defined in advance by the user or automatically determined by the algorithm. Discovering the optimum number of clusters or natural groups in the data is not a trivial task. The popular clustering techniques which are suggested so far are either partition based or hierarchy based, but both approaches have their own advantages and limitations in terms of the number of clusters, shape of clusters, and cluster overlapping[3] .Some other approaches are designed using different clustering techniques and involve optimization in the process. The involvement of intelligent optimization techniques has been found effective to enhance the complex, real time, and costly data mining process.

II. K-Means Algorithm

The conventional K-mean algorithm is based on decomposition, most popular technique in data mining field. The concept of K-Means algorithm uses K as a parameter, Divide n object into K clusters, to create relatively high similarity in the cluster and, relatively low similarity between clusters. And minimize the total distance between the values in each cluster to the cluster center. The cluster center of each cluster is the mean value of the cluster. The calculation of similarity is done by mean value of the cluster objects. The measurement of the similarity for the algorithm selection is done by the reciprocal of Euclidean distance. That is to say, the closer the distance, the bigger the similarity of two objects, and vice versa.

a) Procedure of K-mean Algorithm

K-mean distributes all objects to K number of clusters at random;

1) Calculate the mean value of each cluster, and use this mean value to represent the cluster;
2) Re-distribute the objects to the closest cluster according to its distance to the cluster center;

Author α : Ritu Yadav , Department of Computer Science and Engg,Guru Jambheshwar University of Science & Technology,Hissar,Haryana India. E-mail:ryadav1986@gmail.com
Author σ : Anuradha Sharma,aDepartment of Computer Science and Engineering , Rajasthan Institute of Engineering & Technology, Rajasthan Technical University,Jaipur. Rajasthan,India. E-mail:anuradha.sharma918@gmail.com

3) Update the mean value of the cluster, say, calculate the mean value of the objects in each cluster;
4) Calculate the criterion function E, until the criterion function converges.

Usually, the K-mean algorithm criterion function adopts square error criterion, defined as:

$$E = \sum_{i=1}^{k} \sum_{p \in C_i} |p - m_i|^2$$

In which, E is total square error of all the objects in the data cluster, p is given data object, mi is mean value of cluster Ci (p and m are both multi-dimensional). The function of this criterion is to make the generated cluster be as compacted and independent as possible [4].

b) Analysis of the Performance of the K-mean Algorithm

i. *Advantages*

1) It is a classic algorithm to resolve cluster problems; this algorithm is simple and fast;
2) For large data collection, this algorithm is relatively flexible and highly efficient, because the complexity is O (ntk), among which, n is the number of all objects, k is the number of cluster, t is the times of Iteration. Usually, k<<n and t<<n. The algorithm usually ends with local optimum.
3) It provides relatively good result for convex cluster;
4) Because of the limitation of the Euclidean distance [5]. It can only process the numerical value, with good geometrical and statistic meaning;

ii. *Disadvantages*

1) Sensitive to the selection of initial cluster center, usually end without global optimal solution, but suboptimal solution;
2) There is no applicable evidence for the decision of the value of K (number of cluster to generate), and sensitive to initial value, for different initial value, there may be different clusters generated;
3) This algorithm is easy to be disturbed by abnormal points; a few of this abnormal data will cause extreme influence to the mean value;
4) Sometimes the result of cluster may lose balance

III. Advanced K-Means Clustering Algorithms

The K-means algorithm and its conjoined algorithms are in the family of center base clustering algorithms. This family have several methods: expectation maximization, fuzzy K-means and harmonic K-means. A brief review of these algorithms is given in the following sub-sections.

a) General Clustering Algorithm

D. Mayszko [6], proposed the steps for the clustering algorithm are as follows:

1) Initialize step with centers C.
2) For each data point xi, compute its minimum distance with each center cj.
3) For each center cj, recomputed the new center from all data points xi belong to this cluster.
4) Repeat steps 2 and 3 until convergence.

b) Expectation Maximization

According to C.M. Bishop, Expectation maximization algorithm uses a linear combination of Gaussian distribution as centers [7]. Its minimization is:

$$EM(X,C) = \sum_{i=1}^{n} \log\left(\sum_{j=1}^{k} p(x_i|c_j)p(c_j)\right)$$

This algorithm has a constant weight that gives all data point to its nearest center.

c) Fuzzy K-Means

According to S. Wierzchoń, Fuzzy K-means algorithm is also called fuzzy c-means. It is adaptation of the K-means algorithm and use soft membership function. This algorithm determines a data point belongs to any centers depends on its membership as [6]:

$$FKM(X,C) = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} \|x_i - c_j\|^2$$

This algorithm has a soft membership and constant weight that gives all data point to the closed center.

d) Harmonic K-Means Algorithm

According to B. Zhang, The harmonic K-means algorithm is a method which is similar to the standard K-means. It uses the harmonic mean of the distance from each data point to all centers as [8]:

$$HKM(X,C) = \sum_{i=1}^{n} \frac{k}{\sum_{j=1}^{k} \frac{1}{\|x_i - c_j\|^2}}$$

This algorithm has a soft membership and weight function to points that are far away from every center.

e) Early Stop K-Means

Early stop K-means algorithm is the first one to handle a convergence step in the standard K-means algorithm. It consists of associating the square error values to a convergence condition. It gets action when there are two consecutive iterations and the square error of the last iteration exceeds that of the preceding iteration. It finds a solution at least as good as that of the standard K-means with a number of iterations smaller than or equal to that of standard K-means algorithm [9].

f) Modified K-Means

According to W. Li, Modified K-means algorithm is a new algorithm for K-means based on the

optimization formulation and a novel iterative method. The steps of this algorithm represented as [10]:

1) Dividing data set (D) into K parts:

$$D = \cup_{k=1}^{K} S_k, S_{k1} \cap S_{k2} = \emptyset, k1 \neq k2$$

2) Let $x_{(k)}^{(0)}, k = 1, \ldots, K$ be initial clustering centers calculate by:

$$x_{(k)}^{(0)} = \sum_{d^{(j)} \in S_k} d^{(j)} / |S_k|, \quad k = 1, \ldots, K.$$

3) Decide membership of the patterns in each one of the K clusters according to the minimum distance from cluster center.
4) Calculate new centers using the iteratively.
5) Repeat step 3 and 4 till there in no change in cluster center.

IV. Methods to Improve K-Means Algorithm's Performance

a) *Methods for Initial Point Selection*

i. *Refining Initial Points Algorithm*

In partitional clustering algorithm, the first step we should get initial seed points (cluster centers). To choose good initial points will improve solutions and reduce execution time. Refining initial points algorithm is proposed[2].

For a start, randomly choose some subsets within equal number of samples from large data sets. Secondly, partitional algorithm is applied to each subsets to get each center sets of the subsets. Thirdly, gather these center sets and apply the partitional algorithm again to obtain the most proper center set. For getting good initial seed points, totally we repeat the partitional algorithm 2 times by fewer sample sets. Finally run the partitional algorithm with the most feasible center set as the initial seeds and original large data sets.

The algorithm steps are:

1) Randomly build J sample subsets. Si is a random subset of data (i = 1...J and the size of Si is S_size).
2) Use modified algorithm to find center Ci of each Si . Gather all Ci (i = 1...J) into C _Total .
3) For each set Ci (i = 1...J), run paritional algorithm with initial points Ci and data set C _Total to get another center set FCi .
4) For each FCi(i = 1...J), calculate sum Sumi of the distance between each point in C_Total to the closest center point in FCi .
5) Find minimum of Sumi (i =1...J). If Sump is minimum, take FCp as final initial points.

ii. *Cluster Centroid Decision Method*

This method proposed a technique to assign the data point to appropriate cluster's centroid, we calculate the distance between each cluster's centroid and for each centroid take the minimum distance from the remaining centroid and make it half, denoted by dc(i) i.e. half of the minimum distance from ith cluster's centroid to the remaining cluster's centroid. Now take any data point to calculate its distance from ith centroid and compare it with dc(i). If it is less than or equal to dc(i) then data point is assigned to the ith cluster otherwise calculate the distance from the other centroid. Repeat this process until that data point is assigned to any of the remaining cluster. If data point is not assigned to any of the cluster then the centroid which shows the minimum distance with data point becomes the cluster for that data point. Repeat this process for each data point. Take mean of each cluster separately and update the centroid of clusters like traditional k-mean. Repeat this process until termination condition is achieved[3].

N0: Number of data point
K : Number of cluster's centroid
Ci: ith cluster
Some equations used in algorithm are:

$$|Ci,Cj| = \{d(mi,mj) : (i,j)\varepsilon\ [1,k] \ \&\ i \neq j\}$$

Where $|Ci,Cj|$ is the distance between cluster Ci and Cj.

$$dc(i) = \tfrac{1}{2}(\min\{|Ci,Cj|\}$$

where dc(i) is the half of the minimum distance from ith cluster to any other remaining cluster

iii. *Cluster Seed Selection*

When calculating the K turn of clustering seeds with the improved algorithm, those data in the cluster having a great similarity to the K-1 category seeds should be adopted to calculate their mean points (geometrical center) as the clustering seed of the K tum and the specific calculation method is below as[13]:

1) For the cluster Ci(k-1) obtained through the K-1 tum of clustering, the minimum similarity sim _ mini(K-1) of the data in the cluster to the clustering seed Si(k-l) of the cluster is calculated;
2) The data in the cluster Ci(k-l) is calculated that has a similarity of more than 1-β* (1-sim _ mini(k-l))to the clustering seed Si(k-l) (among, β is a constant between 0-1), and the data set is recorded as cni(k-1)
3) The mean points of the data in cni(k-l) are calculated as the clustering seed of the K tum.

b) *Methods to Define no of Cluster*

i. *Initialization Method*

This method depends on the data and works well to find the best number of cluster and their centroids values. It starts by reading the data as 2D matrix, and then calculates the mean of the first frame size F1=300x300,F2=150x150, F3=100x100, F4=50x50, F5=30x30, F6=10x10 or F7=5x5. Then, it keeps the value of means in an array called means array

even at the end of the data matrix. After that it sorts the values in the means array in an ascending manner. In cases where the values are similar, they are removed to avoid an overlap. In other words, only one value is kept. It will then calculate the number of elements in the means array: this number is the number of clusters and their values are the centroids values as indicated in the steps below as [11]:

1) Read the data set as a matrix.
2) Calculate the means of each frame depending on the frame size and putting them in the means array.
3) Sort the means array in an ascending way.
4) Comparing between the current element and the next element in the means array. If they are equal, then keep the current element and remove the next, otherwise, keep both.
5) Repeat step 4 until the end of the means array.
6) Count how many elements remain in the means array. These are equal to the number of clusters and their values.

ii. *The Encoding Method*

According to the characteristics of K-mean cluster algorithm, to find the optimum cluster, the optimum K value should be found, the value of K is the learning object of the genetic algorithm, the encoding it encoding to K value. In general situation, to the class issue, there is always a maximum number of classes "MAXCLassnum" for the cluster, this value is input by the user. So K is a integral between 1 and MAXCLassnum, can be indicated in a binary string. In this experiment, using a byte to express K value, that is 255 classes maximum. This value is enough for normal cluster problem[4].

1) Chooses n number of chromosome from the original n chromosome Using the roulette wheels selection of the traditional genetic algorithm.
2) Crossover method is applied on selected chromosome in the matting pool.
3) Mutation is applied over chromosomes in the matting pool.
4) Form a new generation of chromosome with the original chromosome.
5) Design the fitness function to evaluate K value by the quality of sample cluster result.

The Fitness function is:

$$Fitness = \omega_1 \frac{Dis\ \ of\ \ class}{1 + Dis\ \ in\ \ class} + \omega_2 \frac{1}{NumDifference}$$

Distance between classes is:

$$Dis\ \ of\ \ class = \frac{2\sum_{i=0}^{k}\sum_{j=i+1}^{k} dis(center_i, center_j)}{k(k-1)}$$

The centeri is classi of cluster center, dis(x,y)is the Euclidean distance between x,y.

Distance between classes is:

$$Dis\ \ in\ \ class = \frac{1}{k}\sum_{i=0}^{k}\left(\frac{\sum_{j=0}^{num_i} dis(Sample_{ij}, center_i)}{num_i}\right)$$

the numi is number of class of i, sample is the samplej of the classi, NumDifference show the statistics of the difference of sample between classes.The ω1, ω1 of Fitness function is the weight of distance between classes.

iii. *Tentative Clustering*

Clustering uses principal components analysis, to determine a tentative value of count of classes and provide changeable labels for objects. The kernel based clustering approach performs principal component analysis on standard score of a given matrix and thereafter projects the matrix into space of the calculated principal vectors. Count of the employed principal vector is depending on the given number of classes (K-Means algorithm needs 'K' to process). In order to avoid depending on the number of classes 'K' and to find maximum possible classes, we project the matrix to the space of all principal vectors. After we calculate a probability matrix (P) from result of projection (matrix C) such that Pi,j entry shows probability of connectivity of ith object to jth object.

By refining matrix C according to the probability values of matrix P, we find a block matrix that represents groups of objects[12].

V. Conclusion

This paper presents an overview of the k-means clustering algorithm. K-means clustering is a common way to define classes of jobs within a Dataset. The initial starting point selection may have a significant effect on the results of the algorithm, both in the number of clusters found and their centroids. Methods to improve performance of k-means clustering are discussed in this paper. These methods fall into two categories: initial point selection and define number of cluster. Six of these methods, three from each category, are presented. These methods have been implemented in data mining system and can get better results for some practical programs such as character recognition, image processing, text searching .

References Références Referencias

1. J. Han and M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, San Francisco, CA, 2001.
2. Jirong Gu; Jieming Zhou; Xianwei Chen;"An Enhancement of K-means Clustering Algorithm", in proceeding of Business Intelligence and Financial Engineering, 2009. BIFE '09. International Conference ,2009, Key Lab. of the Southwestern Land Resources Monitoring, Sichuan Normal Univ., Chengdu, China

3. Singh, R.V.; Bhatia, M.P.S.;" Data Clustering with Modified K-means Algorithm"in proceeding of Recent Trends in Information Technology (ICRTIT), 2011 International Conference,2011, Dept. of Comput. Sci. & Eng., Univ. of Delhi, New Delhi, India
4. Juntao Wang; Xiaolong Su; "An improved K-means Clustering Algorithm", in proceeding of Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference,2011, Sch. of Comput. Sci. & Technol., China Univ. of Min. & Technol., Xuzhou, China
5. Colm Kearney and Andrew J. Patton, "information gain ranking", Financial Review, 41, 2000, 29-48
6. D. Mła yszko, S. T. Wierzchoń "Standard and Genetic K-means Clustering Techniques in Image Segmentation", (CISIM'07) 0-7695-2894-5/07 IEEE 2007
7. C. M. Bishop, " Neural networks for pattern recognition ", Clarendon Press, Oxford, 1995.
8. B. Zhang, "Generalized k-harmonic means - Boosting in unsupervised learning", Technical Report HLP-2000-137", Hewlett-Packard Labs,2000.
9. J. Pérez, R. Pazos , L. Cruz, G. Reyes, R. Basave, and H. Fraire "Improving the fficiency and Efficacy of the K-means Clustering Algorithm Through a New Convergence Condition", Gervasi and M.Gavrilova (Eds.): ICCSA 2007, LNCS 4707, Part III, pp. 674–682. Springer-Verlag Berlin Heidelberg 2007.
10. W. Li "Modified K-means clustering algorithm", 978-0-7695-3119-9/08 , 2008 IEEE, DOI 10.1109/CISP.2008.349
11. Samma, Ali Salem Bin; Salam, Rosalina Abdul;" daptation of K-Means Algorithm for Image Segmentation"in proceeding of International Journal of Signal Processing,2009
12. Dashti, H.T.; Simas, T.; Ribeiro, R.A.; Assadi, A.; Moitinho, A.; "MK-means - Modified K-means clustering algorithm", in proceeding of Neural Networks (IJCNN), The 2010 International Joint Conference,2010, Univ. of Wisconsin, Madison, WI, USA
13. Li Xinwu;" Research on Text Clustering Algorithm Based on Improved K-means",in proceeding of Computer Design and Applications (ICCDA), 2010 International Conference ,2010, Jiangxi Univ. of Finance & Econ., Nanchang, China

This page is intentionally left blank

Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

TCP Performance Over Mobile IP Wired-cum- Wireless Networks

By Md. Asif Nashiry, Shauli Sarmin Sumi, Subrata Kumar Das & Md. Kamrul Islam

Jessore Science & Technology University, Jessore, Bangladesh

Abstract - Reliable transport protocol like TCP has served well the wired Internet where the packet losses are mainly due to congestion, but is not ready for Mobile IP wired-cum-wireless environments where the significant packet losses are due to bit errors and handoffs. In this paper, we have investigated the performance of TCP among the various TCP variants. We have observed both TCP senders (Newreno & Vegas) and TCP receivers (Base & Delayed-Ack). Using ns-2, we have evaluated the TCP throughput and packet delay over a single TCP connection. The simulation results suggest that a particular combination (one TCP sender with one TCP receiver) of TCP shows the best result in such Mobile IP networks.

Keywords : *Wired-cum-Wireless Networks, TCP, Mobility, Mobile IP, Handoff.*

GJCST Classification: *C.2.1,C.2.5*

Strictly as per the compliance and regulations of:

TCP Performance Over Mobile IP Wired-cum-Wireless Networks

Md. Asif Nashiry [α], Shauli Sarmin Sumi [σ], Subrata Kumar Das [ρ] & Md. Kamrul Islam [ω]

Abstract - **Reliable transport protocol like TCP has served well the wired Internet where the packet losses are mainly due to congestion, but is not ready for Mobile IP wired-cum-wireless environments where the significant packet losses are due to bit errors and handoffs. In this paper, we have investigated the performance of TCP among the various TCP variants. We have observed both TCP senders (Newreno & Vegas) and TCP receivers (Base & Delayed-Ack). Using ns-2, we have evaluated the TCP throughput and packet delay over a single TCP connection. The simulation results suggest that a particular combination (one TCP sender with one TCP receiver) of TCP shows the best result in such Mobile IP networks.**

Keywords: Wired-cum-Wireless Networks, TCP, Mobility, Mobile IP, Handoff.

I. Introduction

TCP was basically developed assuming that it would run on wired networks. Wired networks usually have less bit error rates and hence less packet loss. Also, since mobility was not considered, there are no packet losses and delays caused by mobility. Hence, TCP was designed to assume that any segment loss is caused by the network congestion and on seeing any segment loss; TCP invokes its congestion control measures. In various circumstances wired and wireless networks are connected together to take the advantages of both. Usually, the wireless links have high bit error rates. Also, temporary disconnections occur because of the factors like channel fading, and handoffs when a mobile node is in motion. The handoff period depends on both the link level handoff protocol being used as well as the IP level handoff protocol being used, if any. The standard TCP implementations assume the cause of any packet loss to be the network congestion. Then, they reduce the congestion size to a minimum. Also, TCP invokes the slow-start mechanism. If the network links are slow, it takes long time to grow the congestion window. Thus, reduction in the size of the congestion window reduces the transmission rate and hence degrades the performance. Also, TCP undergoes a binary exponential backoff, causing long pauses of communication. Because of this, some time may be unused even after the mobile node recovers from the temporary disconnection. It has been observed that even a single wireless link can reduce the TCP performance considerably. In theory we can use existing transport protocols like TCP on a wireless host to communicate with a fixed network. Though this keeps the transport layer transparent to mobility, disconnection and other features of wireless and mobile hosts reduce the performance. Thus TCP proved not so efficient in Mobile IP wired-cum-wireless environments and suffers from performance degradation introduced by the conditions that exist in those environments. Hence, an optimized reliable transport protocol is very important for the development of the wireless Internet. This paper provides a realistic comparison of the performances of the various TCP variants over Mobile IP wired-cum-wireless network. The TCP variants include both TCP senders (TCP Newreno and TCP Vegas) and TCP receivers (Base TCP Sink and Delayed-Ack TCP Sink). The reminder of this paper is organized as follows: Section 2 depicts the related works; mobile IP and TCP variants are described in Section 3. We define the simulation model in Section 4 and present the results and discussions in Section 5. The conclusion and future work follow in Section 6.

II. Related Works

TCP/IP is the standard networking protocol stack for Internet. So it is expected to be deployed over wireless networks to allow seamless integration with the Internet. However, earlier research suggested that TCP performs poorly over cellular (single-hop) wireless networks (Holland & Vaidya, 1999). A good amount of research has been done in the past to improve the performance of TCP in the conditions that prevail in mobile and wireless environments. Split connection approach is one of them. It separates TCP connection, one for the wireless network and the other for the wired network between a sender and a receiver. The motivation behind the use of separate connections is the fact that the characteristics of a wireless network are different from the characteristics of a wired network, hence a separate connection that is optimized to the

Author α : Md. Asif Nashiry, Department of Computer Science & Engineering, Jessore Science & Technology University, Jessore, Bangladesh. E-mail: asif.nashiry@yahoo.com

Author σ : Shauli Sarmin Sumi, Department of Information & Communication Engineering, Islamic University, Kushtia-7003, Bangladesh. E-mail: shaulisumi@yahoo.com

Author ρ : Subrata Kumar Das, Department of Computer Science & Engineering, Jessore Science & Technology University, Jessore, Bangladesh. E-mail: sdas_ce@yahoo.com

Author ω : Md. Kamrul Islam, Department of Computer Science & Engineering, Jessore Science & Technology University, Jessore, Bangladesh. E-mail: polash_2k4@yahoo.com

kind of network it operates on, can be used. I-TCP (Bakre & Badrinath, 1995), ELN (Balakrishnan & Katz, 1998) can be classified in this group.

Enhanced link layer approach is another solution. Since the basic factor that causes degradation in wireless environments is the nature of the transmission medium, an improvement in the link technology that operates over the physical layer can improve the overall performance. This was the idea behind enhanced link layer solutions. Mechanisms like ARQ, FEC (Chockalingam Zorzi & Tralli, 1999) were suggested to improve the performance at the lower layers which, in turn, can reduce the chances of TCP getting timed out and does not invoke its congestion control measures.

Most of the above-mentioned solutions require special support from the network infrastructure in some form or the other. Some other approaches were suggested that do not require any special support from the intermediate infrastructure and changes are confined to the end nodes. Path Prediction (Hadjiefthymiades, Papayiannis & Merakos, 2002) and ACK Regulator (Chan & Ramjee, 2002) come under this category. These solutions are not bound by the problem of encrypted traffic, and they do not assume any special support from the network infrastructure.

As suggested earlier, most related work focuses on to improve the TCP performance in wireless network. Also most of the earlier solutions (except ELN) try to optimize the mobile host acting as a TCP receiver but do not consider the case of a mobile TCP sender, which can also be a common case in the future. Also a few of the earlier work measured TCP performance with combining both the TCP senders and the TCP receivers. This paper provides a realistic comparison of performance of the various TCP variants over Mobile IP wired-cum-wireless network. The TCP variants include both TCP senders (TCP Newreno and TCP Vegas) and TCP receivers (Base TCP Sink and Delayed-Ack TCP Sink).

III. Mobile IP and TCP Variants

In order to achieve the mobility function, the Internet Protocol (IP) has extended to become the Mobile Internet Protocol (Mobile IP or MIP). Mobile IP provides hosts with the ability to change their point of attachment to the network without compromising their ability to communicate. The mobility support provided by Mobile IP is transparent to the other protocol layers so as not to affect the operation of applications which do not have the mobile capability.

Among various IP mobility proposals, Mobile IPv4 (Charles, 1996) & (Perkins, 2002) is the oldest and probably the most widely known mobility management proposal with IP. MIPv4 introduces three new entities required to support the protocol: the Home Agent (HA), the Foreign Agent (FA) and the Mobile Host (MH). Home and foreign agents are introduced for mobility management. Each time a mobile host connects to a network at a new location, it will obtain a temporary address, called Care-of Address (COA) from a foreign agent in the local network. Then the mobile host must inform its home agent of the new address by a registration procedure, which begins when the mobile host, possibly with the assistance of the foreign agent, sends a registration request with the COA. When the home agent receives this request, it may typically add the necessary information to its routing table, approve the request, and send a registration reply back to the mobile host.

A basic function of TCP is to provide reliable communication over an unreliable network layer. TCP ensures that the data is delivered from the sending process to the receiving process correctly and in order. TCP controls the traffic flow and network congestion by maintaining window (buffer to store data packets); and various TCP variants are classified according to the mechanism for maintaining this window.

a) TCP New-Reno

A modification to Reno TCP called New-Reno TCP (Hoe, 1996). Since most TCP sessions last only for a short period of time, the initial slow start period is significant for the overall performance. A method is proposed to estimate an optimum slow start threshold (ssthresh) value by calculating the byte equivalent of bandwidth delay product of the network when a new connection is made. TCP New-Reno also deals with multiple packet losses from a single window. If two or more segments have been lost from the transmitted data (window), the Fast Retransmission and Fast Recovery algorithms will not be able to recover the losses without waiting for retransmission time out. New-Reno overcomes this problem by introducing the concept of a Fast Retransmission Phase, which starts on detection of a packet loss (receiving 3 duplicate ACKs) and ends when the receiver acknowledges reception of all data transmitted at the start of the Fast Retransmission phase.

The transmitter assumes reception of a partial ACK during the Fast Retransmission phase as an indication that another packet has been lost within the window and retransmits that packet immediately to prevent expiry of the retransmission timer. New Reno sets the congestion window (cwnd) to one segment on reception of 3 duplicate ACKs (i.e. when entering the Fast Retransmission Phase) and unacknowledged data are retransmitted using the Slow Start algorithm. The transmitter is also allowed to transmit a new data packet on receiving 2 duplicate ACKs. While the transmitter is in the Fast Retransmission Phase, it continues to retransmit packets using Slow Start until all packets have been recovered (without starting a new

retransmission phase for partial ACKs). Although this modification may cause unnecessary retransmissions, it reduces transmitter timeouts and efficiently recovers multiple packet loss using partial ACKs.

b) TCP Vegas

TCP Vegas (Brakmo & Peterson, 1995) does not involve any changes to the TCP specification; it is merely an alternative implementation that is able to interoperate with any other valid implementation of TCP. In fact, all the changes pare confined to the sending side. There are some techniques that Vegas employs to increase throughput and decrease losses.

It calculates the RTT more accurately. Using an accurate RTT estimate serves two purposes. First, it leads to a more accurate timeout calculation. Second, it is used in conjunction with the mechanism described next to decide to retransmit a dropped segment in a more timely fashion. It includes the new mechanism for deciding to retransmit Vegas treats the receipt of certain ACKs as a trigger to check if a timeout should happen. Moreover, Vegas only decreases the congestion window if the retransmitted segment was previously sent after the last decrease. Any losses that happened before the last window decrease do not imply that the network is congested for the current congestion window size, and therefore do not imply that it should be decreased again. Another most important feature of TCP Vegas is that the detection of congestion. Vegas calculates the expected throughput by:

Expected = WindowSize / BaseRTT

where WindowSize is the size of the current congestion window and BaseRTT is the RTT of a segment when the connection is not congested. It then calculates the current Actual throughput of the network. Then, Vegas compares Actual to Expected, and adjusts the window accordingly.

c) Base TCP Sink

The base TCP sink object is responsible for returning ACKs to a peer TCP source object. It generates one ACK per packet received. The size of the ACKs may be configured. The creation and configuration of the TCP sink object is generally performed automatically by a library call.

d) Delayed-Ack TCP Sink

A Delayed-Ack sink object is available for simulating a TCP receiver that ACKs less than once per packet received. This object contains a bound variable, which gives the number of seconds to wait between ACKs. The Delayed-Ack sink implements an aggressive ACK policy whereby only ACKs for in-order packets are delayed. Out-of-order packets cause immediate ACK generation. It is possible to configure the bound variable (interval) of the Delayed-Ack TCP Sink.

IV. Simulation Model & Design

To design the Mobile IP wired-cum-wireless network topology we have considered various simulation parameters, which are supported by the network simulator-2 (NS-2) (Fall & Varadhan, 2005). We have done simulation for two different topologies and call them topology 'A' and topology 'B' respectively. There is a roaming mobile node called Mobile Host (MH) that moves between its home agent and foreign agent in a certain velocity. Data will be exchanged

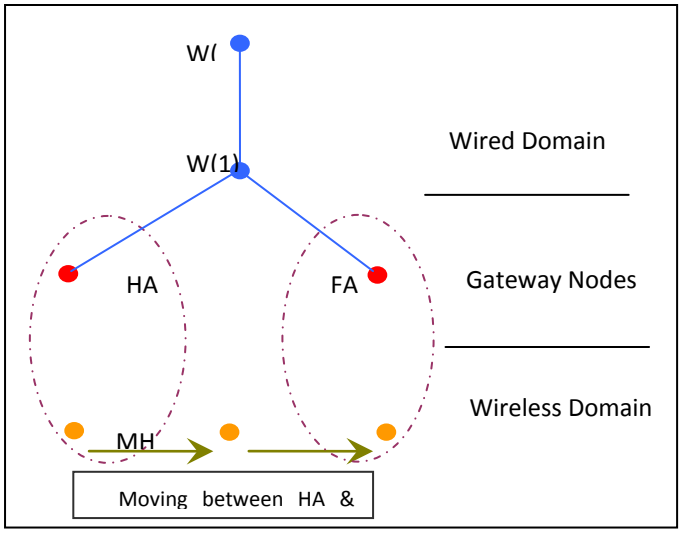


Figure 1: Topology 'A' for Mobile IP wired-cum-wireless simulation

between the roaming mobile node (MH) and the wired node W(0) via the two base-station nodes (HA or FA). Base station nodes are like gateways between wireless and wired domains and allow packets to be exchanged between the two different types of nodes. The initial positions, (x,y), of the nodes MH, HA and FA are

(50,100), (200,300) and (500,300) respectively in the 700mX800m topology. In this network, the mobile node MH starts moving from position (50,100) at 10 sec towards destination (650,100) at fixed speeds (0, 5, 10, 15, 20, 25, or 30) m/s. After reaching the destination position (650,100), MH will stay there 1sec (pause time) and then again moves back to its initial starting position (50,100) at the same velocity. In this way, the mobile node MH changes its home agent (HA) and foreign agent (FA) several times depending upon its speed. We have used DSDV ad hoc routing protocol. FTP begins transferring packets of size 1000 bytes after 10 second. Each wireless node has a buffer size of 50 packets. Each simulation ends at 150 seconds.

There are fewer changes in Topology 'B'. We have used three base-station nodes in topology 'B' and call them Home Agent (HA), Foreign Agent-1 (FA1) and Foreign Agent-2 (FA2) respectively. The initial positions, (x,y), of the nodes MH, HA, FA1 and FA2 are (50,100), (200,300), (500,300) and (800,300) respectively in the 1000mX800m topology. In this case, simulation starts at 10 sec and ends at 190 sec. Here the mobile node MH starts moving from position (50,100) at 10 sec towards destination position (950,100) at a fixed speed of (0, 5, 10, 15, 20, 25, or 30) m/s while nodes HA, FA1 and FA2 serve as base station (gateway) nodes. As before, the pause time is 1 sec, that is, the mobile node MH will stay 1 sec at position (950,100) and then moves back to its initial position (50, 100) at the same velocity. Staying there 1 sec, it moves again towards the destination position.

V. Performance Analysis & Discussion

This section describes a realistic analysis comparing the performance of the various TCP variants over multi-hop Mobile IP wired-cum-wireless networks. We have combined the TCP senders (New-Reno & Vegas) with the TCP receivers (Base & Delayed-Ack). We have considered the mobile node as a TCP sender.

a) Throughput Analysis

We have evaluated the throughput for topology 'A' obtained with the TCP variants in terms of the total number of packets received at the destination node per unit time over intervals of 7 seconds for mobility 0 meter/sec, 15 meter/sec and 30 meter/sec and the results are plotted in figures.

From figure 2, we observe that after the initial connection establishment, all the variants achieve almost fixed throughput to the whole simulation time. It is common that if the window size increases, the throughput will also increase.

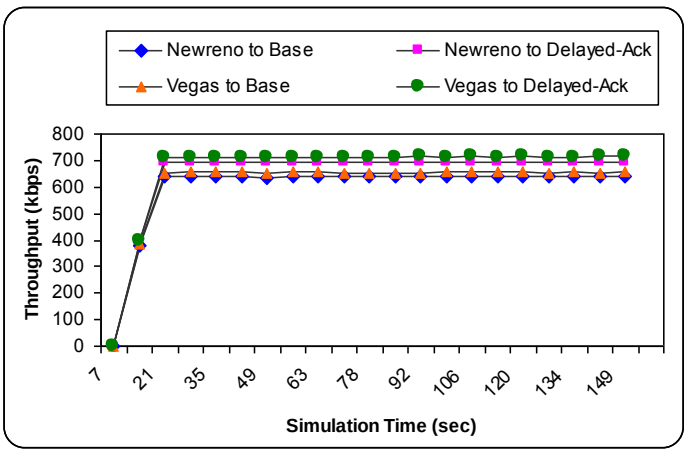


Figure 2 : Throughput (kbps) versus Simulation Time for topology 'A' for mobility 0 (meter/sec)

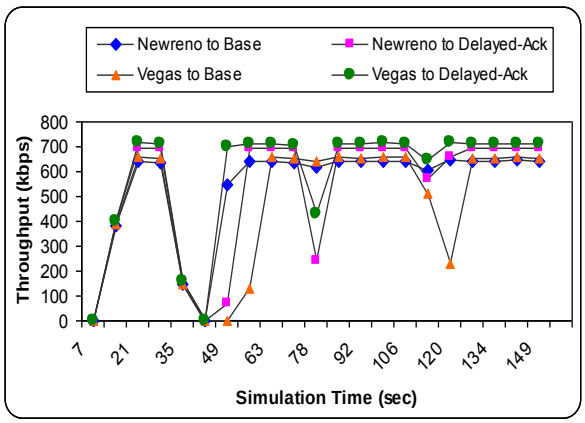


Figure 3(a) : Mobility 15 (m/s)

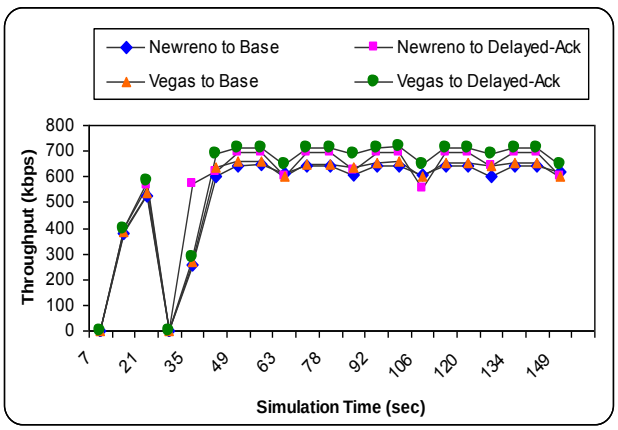


Figure 3(b) : Mobility 30 (m/s)

Figure 3 : Throughput (kbps) versus Simulation Time for topology 'A'

Among the four variants, Vegas to Delayed-Ack shows the best throughput. From figure 3 (a), it is seen that the throughput of each of the TCP variant decreases drastically. By comparing this figure to figure 2, it is observed that the throughput performance is

worse in this case. From the start of the simulation to 35 seconds, throughput is the same (almost 700 kbps) as before (figure 2). But after 35 seconds it shows different result. The mobility of the mobile host is the cause of this degradation. In the case of figure 2, the mobile host remains stationary. But in this case the mobile host changes its base station for a number of times during simulation time. When changing base station a mobile host requires a handoff. The mobile host changes base station three times; so require three handoffs. And at the time of each handoff packets are dropped and throughput decreases.

Figure 3(b) shows the similar type of result except the throughput is reduced more times than that of figure 3(a). Since in this case, the mobility is 30 meter/sec, more handoff is required and at the time of each handoff throughput is reduced. One interesting feature which is observed from figures 3(a) and 3(b) is that during first handoff throughput performance decreases drastically but during the second, third and successive handoffs throughput decreases slightly. But in such an environment Vegas to Delayed-Ack performs well among the TCP variants. Its throughput reduces less than that of the others. Concentrating on the receiver portion in figure 2, 3(a) and 3(b), we see that the performance of the Base TCP receivers is lower than that of the Delayed-Ack TCP receivers; no matter whether the TCP sender is New-Reno or Vegas. Since all the terminals in Mobile IP network share the same radio channel to send and receive packets, the collision of data packets and acknowledgement (ACK) packets in the radio channel can severely reduce TCP throughput. But Delayed-Ack receiver employs less ACK packets. So, for transferring TCP data packets, TCP sender can utilize the channel more. Hence throughput of any TCP sender synchronized with Delayed-Ack receiver is better than that of TCP sender with Base receiver.

Now with fixed Delayed-Ack receiver, we are going to observe the performance of New-Reno and Vegas sender. By evaluating the throughput from three different mobility scenarios it is seen that TCP Vegas sender performs well. As stated earlier, TCP Vegas estimates the network status to adjust its window value. This property enriches its performance in such a heterogeneous environment.

We also average the throughput over the simulation duration and plot the average throughput obtained with the TCP variants as a function of node mobility for both topology 'A' and 'B' in figure 4 and figure 5 respectively.

According to figure 4 and 5, when the source node is stationary (node mobility 0 meter/sec), all the TCP variants achieve their own highest throughputs in both topology 'A' and topology 'B', because of occurring no handoffs in this situation.

The throughputs achieved with New-Reno to Base TCP and Vegas to Base TCP are lower than those achieved with New-Reno to Delayed-Ack TCP and Vegas to Delayed-Ack TCP in both topology 'A' and topology 'B'. It is due to the fact that ACKs and TCP packets share the same channel; so decreasing the ACKs by using the Delayed-Ack receiver instead of Base receiver it is possible to improve the throughput of TCP.

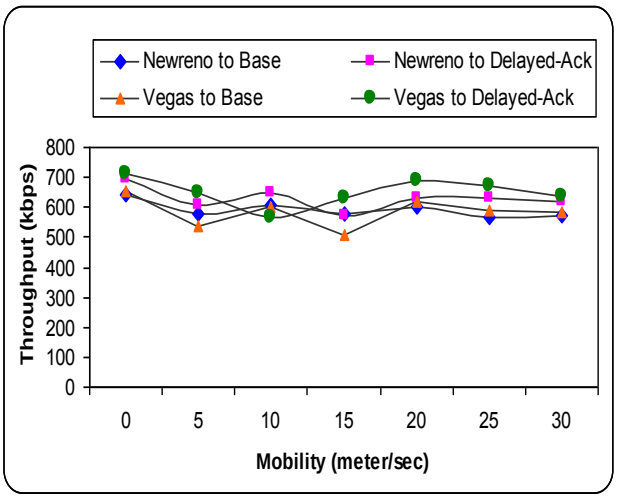


Figure 4 : Throughput vs Mobility for topology 'A'

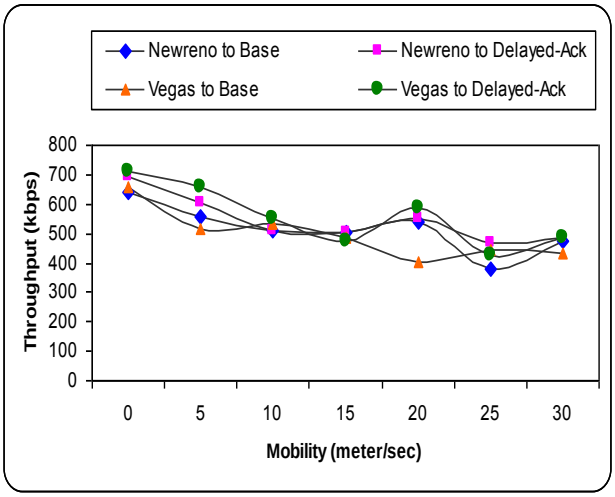


Figure 5 : Throughput vs Mobility for topology 'B'

b) Average Delay Analysis

From figure 6 and figure 7, it is seen that the average delay of New-Reno sender (receiver may be Base or Delayed-Ack) is high as compared to Vegas sender in both topologies. Vegas reads and records the system clock each time a segment is sent. When an ACK arrives, Vegas reads the clock again and does the round trip time (RTT) calculation using this time and a timestamp recorded for the relevant segment. Using an accurate RTT estimates, Vegas leads to a more accurate timeout calculation. This is why Vegas shows very low delay as compared to New-Reno. The average delay of New-Reno to Delayed-Ack is the worst among all the variants. This is why it has given the lowest packet delivery ratio. The average delay of New-Reno to Base is

better than that of New-Reno to Delayed-Ack but worse than Vegas to Base and Vegas to Delayed-Ack. The same thing has happened in case of packet delivery ratio, that is, the packet delivery ratio of New-Reno to Base is better than New-Reno to Delayed-Ack but worse than the Vegas to Base and Vegas to Delayed-Ack.

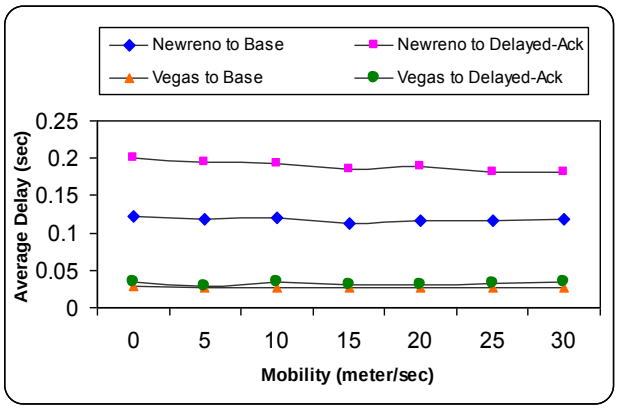


Figure 6 : Average delay vs Mobility for topology A

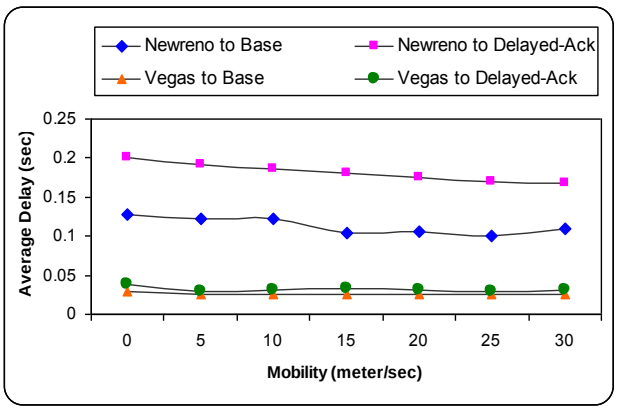


Figure 7: Average delay vs Mobility for topology B

The delay performance of Vegas to Delayed-Ack TCP and Vegas to Base TCP is relatively similar and they show low delay in all the mobility so they show relatively similar and high packet delivery ratio in both the topologies. Since the average delay of Vegas to Delayed-Ack TCP is low it has given a very good throughput in the simulation time in both topologies.

VI. Conclusion and Future Works

While TCP is needed for a Mobile IP network, the current variants of TCP are not adequate for the task. To achieve good TCP performance in Mobile IP wired-cum-wireless networks a variety of solutions have been proposed and in most of the solutions the mobile node acts as a TCP receiver. However, little attention has been paid to study the performance of TCP traffic over combining both TCP senders and receivers. We have investigated the performance of TCP over combining both TCP senders (Newreno & Vegas) and TCP receivers (Base & Delayed-Ack) using simulations in ns-2 for a range of node mobility with a single traffic source. The performance metrics that we have considered include TCP throughput and packet delay. On throughput performance, Vegas TCP to Delayed-Ack gains the best performance. On the sender side, the effective congestion control technique of Vegas makes it optimal in such a heterogeneous environment. And in the receiver side, the Delayed-Ack receiver makes the proper use of channel by sending fewer ACK packets.

We have also evaluated throughput with respect to node mobility. The result shows that throughput of each TCP variant varies similarly with respect to each other against mobility of node. But Vegas to Delayed-Ack shows the best result as compared with other variants. Delay performance of each variant shows the confidential result. TCP Vegas has the lowest delay no matter whether the receiver is Base or Delayed-Ack. New-Reno has large delay. But among them New-Reno to Delayed-Ack has the highest delay. After observing packet delivery ratio and delay, it is seen that the variant that gains highest packet delivery ratio has lowest delay. Though all the variants of TCP suffer in Mobile IP wired-cum-wireless networks, it is clear from the various results that if we use Vegas TCP as a source (sender) and Delayed-Ack as a sink (receiver), then this combination shows the best result among the all TCP variants. Next, we plan to investigate the performance of all the TCP variants with multiple traffic sources in more complex heterogeneous networks and also consider some link layer and network layer feedbacks to improve their performance.

References Références Referencias

1. G. Holland and N. Vaidya (1999). Analysis of tcp performance over mobile ad hoc networks. Proceedings of ACM Mobicom, pp. 219–230.
2. Bakre, A., Badrinath B.R. (1995) .I-TCP: Indirect TCP for Mobile Hosts. Proceedings of the 15th International Conference on Distributed Computing Systems (ICDCS), p. 136-143.
3. Balakrishnan, H., Katz, R. (1998). Explicit Loss Notification and Wireless Web Performance. Proceedings of IEEE Globecom Internet Mini-Conference, Sydney, Australia.
4. Chockalingam, A., Zorzi, A., Tralli, V. (1999). Wireless TCP Performance with Link Layer FEC/ARQ. Proceedings of IEEE ICC'99, p.1212-1216.
5. S. Hadjiefthymiades, S. Papayiannis, and L. Merakos (2002). Using Path Prediction to Improve TCP Performance in Wireless/Mobile Communications. IEEE Communication Magazine.
6. M. C. Chan and R. Ramjee (2002). TCP/IP Performance over 3G Wireless Links with Rate and Delay variation. In Proceedings of MOBICOM.
7. Perkins E. Charles (1996). IP Mobility Support. RFC 2002.

8. Perkins, C (2002), IP mobility support for IPv4, IETF RFC 3344.
9. J. Hoe (1996). Improving the start-up behavior of a congestion control scheme for TCP. In ACM SIGCOMM.
10. L. S. Brakmo and L. L. Peterson (1995). TCP Vegas: End to End Congestion Avoidance on a Global Internet. IEEE Journal on Selected areas in Communications, vol. 13, no. 8, pp. 1465–1480.
11. K. Fall and K. Varadhan (2005) "ns Notes and Documentation. Lawrence Berkeley National Laboratory, 28, http:/www- mash. cs. Berkeley. edu/ns/.

Global Journals Inc. (US) Guidelines Handbook 2012

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

- 'FARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'FARSC" can be added to name in the following manner. eg. **Dr. John E. Hall, Ph.D., FARSC or William Walldroff Ph. D., M.S., FARSC**
- Being FARSC is a respectful honor. It authenticates your research activities. After becoming FARSC, you can use 'FARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 60% Discount will be provided to FARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- FARSC will be given a renowned, secure, free professional email address with 100 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- FARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 15% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- Eg. If we had taken 420 USD from author, we can send 63 USD to your account.
- FARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- After you are FARSC. You can send us scanned copy of all of your documents. We will verify, grade and certify them within a month. It will be based on your academic records, quality of research papers published by you, and 50 more criteria. This is beneficial for your job interviews as recruiting organization need not just rely on you for authenticity and your unknown qualities, you would have authentic ranks of all of your documents. Our scale is unique worldwide.
- FARSC member can proceed to get benefits of free research podcasting in Global Research Radio with their research documents, slides and online movies.
- After your publication anywhere in the world, you can upload you research paper with your recorded voice or you can use our professional RJs to record your paper their voice. We can also stream your conference videos and display your slides online.
- FARSC will be eligible for free application of Standardization of their Researches by Open Scientific Standards. Standardization is next step and level after publishing in a journal. A team of research and professional will work with you to take your research to its next level, which is worldwide open standardization.

- FARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), FARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 80% of its earning by Global Journals Inc. (US) will be transferred to FARSC member's bank account after certain threshold balance. There is no time limit for collection. FARSC member can decide its price and we can help in decision.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

- 'MARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'MARSC" can be added to name in the following manner. eg. Dr. John E. Hall, Ph.D., MARSC or William Walldroff Ph. D., M.S., MARSC
- Being MARSC is a respectful honor. It authenticates your research activities. After becoming MARSC, you can use 'MARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 40% Discount will be provided to MARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- MARSC will be given a renowned, secure, free professional email address with 30 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- MARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 10% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- MARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- MARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), MARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 40% of its earning by Global Journals Inc. (US) will be transferred to MARSC member's bank account after certain threshold balance. There is no time limit for collection. MARSC member can decide its price and we can help in decision.

AUXILIARY MEMBERSHIPS

ANNUAL MEMBER

- Annual Member will be authorized to receive e-Journal GJMBR for one year (subscription for one year).
- The member will be allotted free 1 GB Web-space along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted free 500 MB email space.

PAPER PUBLICATION

- The members can publish paper once. The paper will be sent to two-peer reviewer. The paper will be published after the acceptance of peer reviewers and Editorial Board.

Process of submission of Research Paper

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (*.DOC,*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.Online Submission: There are three ways to submit your paper:

(A) (I) First, register yourself using top right corner of Home page then Login. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal.

(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.

(B) If you are using Internet Explorer, then Direct Submission through Homepage is also available.

(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org.

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11'"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Swis 721 Lt BT.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, "Abstract" word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be three lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

You can use your own standard format also.
Author Guidelines:

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.

2) Drafting the paper and revising it critically regarding important academic content.

3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.

Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.

To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads:

Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications

Research letters: The letters are small and concise comments on previously published matters.

5.STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also.Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

Papers: These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

(a)Title should be relevant and commensurate with the theme of the paper.

(b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.

(c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.

(d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.

(e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.

(f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;

(g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.

(h) Brief Acknowledgements.

(i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and to make suggestions to improve briefness.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

Format

Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than 1.4 × 10-3 m3, or 4 mm somewhat than 4 × 10-3 m. Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:

Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Key Words

A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art.A few tips for deciding as strategically as possible about keyword search:

- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

Acknowledgements: Please make these as concise as possible.

References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

Tables, Figures and Figure Legends

Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.

Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.

Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.

Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.

6. AFTER ACCEPTANCE

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

6.1 Proof Corrections

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

www.adobe.com/products/acrobat/readstep2.html. This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at dean@globaljournals.org within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

6.3 Author Services

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

6.4 Author Material Archive Policy

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

6.5 Offprint and Extra Copies

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: editor@globaljournals.org .

the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

2. Evaluators are human: First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

3. Think Like Evaluators: If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

4. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

5. Ask your Guides: If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

6. Use of computer is recommended: As you are doing research in the field of Computer Science, then this point is quite obvious.

7. Use right software: Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

8. Use the Internet for help: An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

9. Use and get big pictures: Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

10. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

11. Revise what you wrote: When you write anything, always read it, summarize it and then finalize it.

12. Make all efforts: Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

13. Have backups: When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

14. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

15. Use of direct quotes: When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

16. Use proper verb tense: Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

17. Never use online paper: If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

18. **Pick a good study spot:** To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

19. Know what you know: Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

20. Use good quality grammar: Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

21. Arrangement of information: Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

22. Never start in last minute: Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

23. Multitasking in research is not good: Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

24. Never copy others' work: Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

25. Take proper rest and food: No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

26. Go for seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

27. Refresh your mind after intervals: Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

28. Make colleagues: Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

29. **Think technically:** Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

30. Think and then print: When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be

sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

33. Report concluded results: Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

34. After conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.

Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page

- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

· Use standard writing style including articles ("a", "the," etc.)

· Keep on paying attention on the research topic of the paper

· Use paragraphs to split each significant point (excluding for the abstract)

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

· Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives

· Shun use of extra pictures - include only those figures essential to presenting results

Title Page:

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.

Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript--must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to

shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.
- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic

principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

Methods:

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

Approach:

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

What to keep away from

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.

Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.

- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.

Administration Rules Listed Before Submitting Your Research Paper to Global Journals Inc. (US)

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

Segment Draft and Final Research Paper: You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptive of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.

- Do not give permission to anyone else to "PROOFREAD" your manuscript.

- Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.

Criterion for Grading a Research Paper (Compilation) by Global Journals Inc. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

INDEX

A

Achievable · 61, 65, 66, 67
Achieved · 23, 43, 62, 93, 106, 107
Agriculture · 55
Algorithm · 1, 3, 5, 6, 7, 8, 9, 11, 12, 18, 27, 32, 33, 39, 40, 41, 42, 43, 44, 48, 50, 51, 53, 60, 64, 65, 66, 67, 73, 74, 82, 83, 84, 85, 88, 90, 91, 92, 93, 94, 95, 96, 101
Analyzing · 44, 78, 80, 82, 84, 86, 87
Approach · 2, 3, 4, 5, 9, 18, 24, 34, 36, 39, 43, 83, 95, 99, 100
Architecture · 18, 61, 71, 75
Authentication · 32, 41, 71, 72, 73, 74, 75, 77
Automated · 46, 48, 50, 51, 53, 55, 56, 57

B

Beaforming · 58
Benchmarking · 84
Bhattacharjee · 56

C

Candidate · 16, 19, 21, 30, 31
Centroid · 88, 92, 93
Changeable · 95
Characteristics · 58, 60, 61, 62, 64, 66, 69
Chockalingam · 100, 109
Classification · 4, 16, 46, 48, 50, 78, 81, 83, 84, 85, 88
Commitment · 51
Communication · 27, 33, 36, 39, 41, 60, 65, 66, 98, 101
Components · 27, 95
Comprehensive · 71, 73, 75, 77
Computational · 39, 82
Computerized · 36
Concentration · 34
Consecutive · 42, 91
Considered · 51, 52, 53, 58, 60, 61, 65, 71, 74, 98, 102, 104, 109
Convergence · 91
Cryptography · 27, 39, 74, 75
Cybernetics · 12

D

Decomposition · 60, 69, 88
Degradation · 32, 98, 100, 106
Demonstrated · 3, 32, 78, 85
Demonstrates · 64
Denotes · 6
Distribution · 3, 36, 52, 61, 66, 80, 81, 82, 83, 84, 85, 91
Documentation · 110

E

Encryption · 39, 41, 42, 43, 44, 73, 75
Encyclopedia · 37
Equivalence · 79
Exponential · 66
Exportable · 46, 48, 50, 51, 53, 55, 56, 57

F

Featured · 16, 18, 20
Filtering · 16, 18, 19, 20, 21, 22, 23, 24, 26

G

Geometrical · 90, 93

H

Humanware · 35

I

Illustration · 50, 85
Impediments · 55, 60
Implementation · 48, 102
Incorporating · 1, 5, 6, 11
Indispensable · 3, 46
Infrastructure · 46, 56, 74, 100
Intelligent · 12, 13, 46, 69
Irreversible. · 41

J

Jamming · 58, 60, 61, 62, 64, 66, 69

L

Luenberger · 86

M

Mallow · 78, 80, 82, 84, 86, 87
MAXCLassnum · 94
Mechanism · 65, 98, 101, 102
Mobility · 98, 105, 107, 108, 109
Multidimensional · 78, 83, 86

O

Optimization · 48, 54, 60, 62, 78, 81, 82, 88, 92
Optimization · 58, 60, 62, 64, 66, 69
Outperforms · 23, 67

P

Partitioned · 18
Patterns · 1, 3, 5, 7, 9, 11, 12, 13, 15
Performance · 9, 64, 75, 88, 90, 92, 94, 96, 97, 98, 100, 102, 104, 106, 108, 109, 110
Piecewise · 24
Polynomial · 29, 30, 31
Polytechnic · 34, 36, 37, 38
Prefixspan · 1, 6, 9, 11
Presbyterian · 53
Protocol · 39, 41, 43, 44, 69, 75, 100

R

Reasonably · 85
Recommendation · 1, 3, 5, 7, 9, 11, 12, 13, 15
Representation · 52, 73, 86
Represented · 1, 5, 31, 54, 92
Respectively · 43, 64, 80, 102, 104, 106
Restrictions · 58, 59, 67
Retransmission · 101, 102

S

Segmentation · 16, 17, 18, 20, 21, 22, 23, 24, 78, 86
Segmentation · 16, 18, 20, 21, 22, 23, 24, 26, 96
Sequential · 1, 2, 5, 6, 7, 8, 9, 11, 12
Significant · 4, 5, 7, 32, 41, 46, 95, 98, 101
Steganography · 27, 72, 75
Stenographic · 73

T

Techniques · 1, 4, 16, 17, 18, 27, 29, 39, 41, 48, 58, 60, 88, 102
Techniques · 13, 41, 95, 96
Threshold · 7, 8, 58, 101
Throughput · 44, 59, 60, 61, 62, 65, 66, 98, 102, 104, 105, 106, 107, 108, 109
Throughput · 61, 65, 66, 67, 104, 105, 107
Tiruchirappalli · 75
Topologies · 102, 107, 108
Transactions · 71, 73, 75, 77
Transmission · 16, 27, 43, 58, 60, 61, 62, 64, 67, 98, 100
Transmission · 58, 59, 60, 61, 62, 64, 65, 66, 67, 69
Transmitter · 58, 59, 60, 101, 102
Transparency · 21, 23

U

Utilization · 61, 66

W

Watermarking · 27, 29, 31, 32, 33, 75
Weighted · 1, 3, 5, 6, 7, 9, 11, 13, 15, 66
Wireless mesh · 69

save our planet

Global Journal of Computer Science and Technology

Visit us on the Web at www.GlobalJournals.org | www.ComputerResearch.org
or email us at helpdesk@globaljournals.org

ISSN 9754350